（鄂）新登字第 10 号

图书在版编目(CIP)数据

非线性分析·理论与方法/胡适耕 著

武汉：华中理工大学出版社，1996 年 1 月

ISBN 7-5609-1206-0

Ⅰ. 非…

Ⅱ. 胡…

Ⅲ. 非线性-泛函分析

Ⅳ. O177.91

·研究生用书·

非线性分析·理论与方法

胡适耕 著

责任编辑：李立鹏

*

华中理工大学出版社出版发行

（武昌喻家山　　邮编：430074）

新华书店湖北发行所经销

华中理工大学出版社照排室排版

华中理工大学出版社印刷厂印刷

*

开本：850×1168 1/32　印张：9.625　插页：2　字数：240 000

1996 年 1 月第 1 版 1996 年 1 月第 1 次印刷

印数：1-1 500

ISBN 7-5609-1206-0/O·139

定价：7.80 元

## 内 容 简 介

本书系统介绍现代非线性分析的理论、方法与应用. 全书共8章:1. 度理论与不动点;2. 锥与正不动点;3. 单调算子;4. 非光滑分析;5. 非线性最优化;6. 变分不等式;7. 临界点理论;8. 非线性动力系统. 本书一方面以紧凑的形式概括了非线性泛函分析的标准内容,同时较多地介绍了一些近年来日趋重要的新的非线性分析方法,以适应现代数学与自然科学发展的需要. 本书一方面着力改进了基础理论的表述,使之更适于对初学者提供有效导引,同时以适当的篇幅介绍了某些新结果与研究动向,因而为有研究兴趣的读者架设起从基础理论通向研究前沿的桥梁. 本书熔理论、方法与应用于一炉,可适应多方面读者的需要. 对于数学系的高年级大学生及有关理工科专业的硕士生,本书略加删节之后可作为教材使用. 在当代科学发展进程中,对于非线性分析的日益广泛与紧迫的需要,已成为一种引人注目的潮流;有这种需要的科学技术工作者,将发现本书可提供一些有用的理论工具.

## Abstract

The purpose of this book is to provide a systematical treatment for the basic theory, methods and some applications of modern nonlinear analysis. The book includes eight chapters. The main topics are as follows: degree and fixed point theorems; cones, cone mappings and positive fixed points; monotone operators; convex and nonsmooth analysis; nonlinear optimization; variational inequalities; critical point theory and nonlinear dynamical systems, etc. On the one hand the book contains a comprehensive treatment for standard results of nonlinear functional analysis, on the other hand it provides some new concepts and methods of nonlinear analysis which are the subject of great importance and much activity recently. The book can serve as textbook or reference for graduate students. It can also be consulted by relevent teachers, natural scientists and engineers.

# “研究生用书”总序

研究生教材建设是提高研究生教学质量的重要环节，是具有战略性的基本建设。各门课程必须有高质量的教材，才能使学生通过学习掌握各门学科的坚实的基础理论和系统的专门知识，为从事科学研究工作或独立担负专门技术工作打下良好的基础。

我校各专业自 1978 年招收研究生以来，组织了一批学术水平较高，教学经验丰富的教师，先后编写了公共课、学位课所需的多种教材和教学用书。有的教材和教学用书已正式出版发行，更多则采用讲义的形式逐年印发。这些讲义经过任课教师多年教学实践，不断修改、补充、完善，已达到出书的要求。因此，我校决定出版“研究生用书”，以满足本校各专业研究生教学需要，并与校外单位交流，征求有关专家学者和读者的意见，以促进我校研究生教材建设工作，提高教学质量。

“研究生用书”以公共课和若干门学位课教材为主，还有教学参考书和学术专著，涉及的面较广，数量较多，准备在今后数年内分批出版。编写“研究生用书”总的要求是从研究生的教学需要出发，根据各门课程在教学过程中的地位和作用，在内容上求新、求深、求精，每本教材均应包括本门课程的基本内容 ，使学生能掌握必需的基础理论和专门知识；学位课教材还应接触该学科的发展前沿，反映国内外的最新研究成果，以适应目前科学技术知识更新很快的形势；学术专著则应充分反映作者的科

研硕果和学术水平，阐述自己的学术见解。在结构和阐述方法上，应条理清楚，论证严谨，文字简炼，符合人们的认识规律。总之，要力求使“研究生用书”具有科学性、系统性和先进性。

我们的主观愿望虽然希望“研究生用书”的质量尽可能高一些，但由于研究生的培养工作为时尚短，水平和经验都不够，其中缺点、错误在所难免，尚望校内外专家学者及读者不吝指教，我们将非常感谢。

华中理工大学研究生院院长

陳　珽　黄樹槐

1989.11.

# 写在 1995 年

在今天，国家之间的竞争是国家综合实力的竞争，国家综合实力的竞争关键是经济实力的竞争，而经济实力的竞争关键又在于科技（特别是高科技）的竞争，科技（特别是高科技）的竞争归根结底是人才（特别是高层次人才）的竞争，而人才（特别是高层次人才）的竞争基础又在于教育。“百年大计，教育为本；国家兴亡，人才为基。”十六个字、四句话，确是极其深刻的论断。

显然，作为高层次人才培养的研究生教育就在一个国家的方方面面的工作中，占有十分重要的战略地位。可以说，没有研究生教育，就没有伟威雄壮的科技局面，就没有国家的强大实力，就没有国家在国际上的位置，就会挨打，就会受压，就会被淘汰。

“工欲善其事，必先利其器。”教学用书是教学的重要基本工具与条件。这是所有从事教育的专家所熟知的事实。所以，正如许多专家所知，也正是原来的《“研究生用书”总序》中所指出，研究生教材建设是保证与提高研究生教学质量的重要环节，是一项具有战略性的基本建设。没有研究生的质量，就没有研究生教育的一切。

我校从 1978 年招收研究生以来，即着力从事于研究生教材与教学用书的建设。积十多年建设与实践的经验，我校从 1989 年起，正式分批出版“研究生用书”。第一任

研究生院院长陈珽教授就为之写了《"研究生用书"总序》，表达了我校编写这套用书的指导思想与具体要求，"要力求'研究生用书'具备科学性、系统性、先进性"。第二任研究生院长，也就是当时我校的校长黄树槐教授完全赞同这一指导思想与具体要求，从多方面对这套用书加以关心与支持。

我是十分支持出版"研究生用书"的。早在1988年我在为列入这套书中的第一本，即《机械工程测试·信息·信号分析》写"代序"时就提出：一个研究生应该博览群书，博采百家，思路开阔，有所创见。但这不等于他在一切方面均能如此，有所不为才能有所为。如果一个研究生的主要兴趣与工作不在"这一特定方面"，他也可以选择一本有关的书作为了解与学习这方面专业知识的参考；如果一个研究生的主要兴趣在"这一特定方面"，他更应选择一本有关的书作为主要学习用书，寻觅主要学习线索，并缘此展开，博览群书。这就是我赞成为研究生编写系列教学用书的原因。

目前，这套用书已出版了6批共30种，初步形成规模，逐渐为更多读者所认可。在已出版的书中，有8种分获国家级、部省级图书奖，有13种一再重印，久销不衰，有的印刷总数已近万册。采用此套书的一些兄弟院校教师纷纷来信，赞誉此书为研究生培养与学科建设作出了贡献，解决了他们的"燃眉之急"。我们感谢这些赞誉与鼓励，并将这些作为对我们的鞭策与鼓励，"衷心藏之，何日忘之?!"

现在，正是江南初春，"最是一年春好处"。华工园内，

红梅怒放，迎春盛开，柳枝抽绿，梧叶含苞，松柏青翠，樟桂换新，如同我们的国家正在迅猛发展，欣欣向荣一样，一派盎然生机。尽管春天还有乍寒时候，我们国家在前进中还有种种困难与险阻，有的还很严峻，但是，潮流是不可阻挡的，春意会越来越浓，国家发展会越来越好。我们教师所编的、所著的、所编著的这套教学用书，也会在解决前进中的种种问题中继续发展。然而，我们十分明白，这套书尽管饱含了我们教师的辛勤的长期的教学与科研工作的劳动结晶，作为教学用书百花园中的一丛鲜花正在怒放，然而总会有这种或那种的不妥、错误与不足，我衷心希望在这美好的春日，广大的专家与读者，不吝拨冗相助，对这套教学用书提出批评建议，予以指教启迪，为这丛鲜花除害灭病，抗风防寒，以进一步提高质量，提高水平，更上一层楼，我们不胜感激。我们深知，“一个篱笆三个桩”，没有专家的指导与支持，没有读者的关心与帮助，也就没有这套教学用书的今天。

诗云：“嘤其鸣矣，求其友声。”这是我们的心声。

中国科学院院士
华中理工大学校长
兼研究生院院长

杨叔子
于华工园内
1995 年 3 月 7 日

# 前　言

宇宙之内，万物纷呈．或生，或灭，或盛，或衰，或一贯古今，或瞬息万变，莫可逆料．星击长空，山崩海啸，蔚为壮观；阴阳交错，水火相煿，奥妙无穷．面对如此奇妙多姿的世界，艺术家发出古今不绝的讴歌与咏叹，研究者则决心穷究宇宙之法则．运动永恒，无始无终，人们深信冥冥之中自有规律可寻，但追踪之际却难尽得其奥妙！人们只是慨叹世界过于复杂，而数学家则一言以蔽之，曰“非线性而已”．

日月星辰的运行是非线性的，草木虫鱼的繁衍是非线性的，股市的起跌，经济的升降亦是非线性的，….广而言之，世界本质上就是非线性的！研究非线性现象岂能不用非线性的理论与方法！于是乎，各个领域争相“非线性化”起来；于是乎，我们有了非线性力学，非线性光学，…，自然，更少不了非线性数学．再小而言之，就到本书的专题“非线性分析”了．

何谓“非线性分析”？若望文生义，说是“分析非线性问题之谓也”，未免过于宽泛．若界定为“向非线性科学提供数学分析工具之学科”，则似已贴近，但仍不甚确切．著者浅陋，竟不能开宗明义回答这一问题，于是建议读者去翻翻世界性刊物《Nonlinear Analysis：Theory，Methods and Applications》，其中固然绝无“非线性分析”之定义，但信其足以显示出非线性分析在研究什么．本书居然就用上述刊名的一部分“非线性分析・理论与方法”作为书名，“趋同”之意自有，“附会”之心则无．若以本书之区区比拟《Nonlinear Analysis》之浩瀚，那就荒唐了．

本书前两章叙述任意 Banach 空间与有序 Banach 空间中的度理论及由此派生的不动点理论．这部分属于“非线性泛函分析”的传统内容，今天似乎已高度标准化，要作实质性的改进是很难的．

但本书采用了有点"简易"的表述方法,因而只占用较小的篇幅.第三章是关于单调算子理论的一个简要处理,作为该章中心内容的"满射定理",本书采用了尽可能概括的形式.第四章是对"非光滑分析"的一个简要但似不失相对完整的介绍,其内容固然不及专题论述全面,但本书是将"非光滑分析"置于"非线性分析"这一统一框架之内,因而可充分利用前几章的概念、思路与结论,这就在一定程度上简化了处理过程.第五章基本上属于"无限维最优化"这一颇新的课题.无限维最优化对于最优控制、最佳逼近等问题有重要应用,而且也为非线性分析本身提供许多极有价值的思想与方法,其重要性是不容置疑的.短短的第六章立足于前几章所确立的理论框架之内,从单调算子与凸最优化理论演绎出变分不等式的若干基本结果,借以阐明抽象方法在这一领域所能起的作用.关于临界点理论由于已有张恭庆的精采专著[204]可读,本书第七章似乎无足轻重.但如读者希望用较少的时间一睹临界点理论的梗概,或许会感到本书仍有某些方便之处.最后一章涉足非线性动力系统理论.目前,有关浑沌等课题的种种报道有演绎为传奇之势,这无疑大大刺激了各界人士对动力系统理论的兴趣.著者在听了某些"浑沌动力学"讲演之后完全目瞪口呆,且顿生疑虑:数学家将动力系统讲清楚了吗?幸而已有廖山涛[116]、张筑生[209]等优秀著作可读.对于动力系统的标准内容,本书除了处理得更紧凑些之外,别无所为.颇值一提的是,本书独辟一节专论不无重要的"单调动力系统".这曾是著者研究兴趣所在,似应有较大发挥余地,但顾及全书的均衡,只能强制自己搁笔.

用此区区20余万言,辗转游移于多个论题之间,还试图在所涉足的每个方向达到有一定深度的结果,且要使这些结果在某种统一的框架内叶蔓枝连,融为一体——这样的目标能达到吗?著者搁笔之后,仍不胜惶然.或许本来就应抱定一个低得多的目标:本书倘对感兴趣的读者可作引玉之砖,即足以自慰.

本书初稿承蒙何猛省教授详细审阅,许多章节在内容与形式

上的改进都大大得益于何猛省教授的意见.对此,著者谨致以衷心的感谢.著者还要感谢华中理工大学研究生院的支持,没有这种支持,本书根本无法问世.著者也衷心感谢华中理工大学出版社领导及编辑为本书顺利出版而作的大量努力.

作者

1995.1.11

# 记号与约定

$B_r(a)=B(a,r)$：球；$\overline{B}_r(a)=\overline{B}(a,r)$：闭球.

BVP=边值问题.

$C(A,B)$：从 $A$ 到 $B$ 的连续映射之全体.

$C^{1-0}$=局部 Lipschitz 函数类.

$C_c^r$=具紧支集的 $C^r$ 函数类.

$\mathrm{CL}(X)$：$X$ 上的紧线性算子之全体.

$\mathrm{co}A$：$A$ 的凸包；$\overline{\mathrm{co}}A=\overline{\mathrm{co}A}$.

$\mathrm{cone}A$：$A$ 生成的锥；$\overline{\mathrm{cone}}A=\overline{\mathrm{cone}A}$.

$D(F)$：$F$ 的定义域.

$D_f$：$f$ 的有效域.

$Df$：$f$ 的 $G$-导数.

$Df(x,h)$：$f$ 在 $x$ 沿 $h$ 的方向导数.

$D_{\pm}f(x,h)$：单侧方向导数.

$\mathrm{Diff}^r(M)$：$M$ 上的 $C^r$ 同胚之全体.

$d(A,B)$：$A$ 与 $B$ 间的距离.

$d_D(x)=d(x,D)$.

$df$：切映射；$df_x=df(x)$.

$\Delta f(x,h)=f(x+h)-f(x)$.

$\delta_A$：$A$ 的指示函数.

$\mathrm{epi}f$：$f$ 的上图.

$\mathrm{Fix}f$：$f$ 的不动点之全体.

$f'$：$f$ 的 Frechet 导数.

$\mathrm{GL}(X)$：$X$ 的拓扑自同构之全体.

$\mathrm{Gr}F$：$F$ 之图形.

$\mathrm{Homeo}(A)$：$A$ 的自同胚之全体.

$I=\mathrm{id}$：单位映射.

IVP＝初值问题.

$J=[0,1]$(除少数指明的例外).

$\mathscr{K}$＝紧映射类.

Lip＝Lipschitz 映射类.

Lip$f$:$f$ 的 Lipschitz 模数.

$L(X,Y)$:从 $X$ 到 $Y$ 的连续线性算子之全体.

$L(X)=L(X,X)$.

lsc＝下半连续.

ODE＝常微分方程.

PDE＝偏微分方程.

$\mathrm{R}_+=[0,\infty)$;$\mathrm{R}_+^n=(\mathrm{R}_+)^n$,$\mathrm{R}_-^n=-\mathrm{R}_+^n$.

$\overline{\mathrm{R}}=\mathrm{R}\cup\{\infty\}$.

$R(F)$:$F$ 的值域.

$r_\sigma(A)$:$A$ 的谱半径.

SC＝集压缩映射类.

$S^n$:$n$ 维单位球面.

$S_A$:$A$ 的支承函数.

usc＝上半连续.

wlsc＝弱下半连续.

wusc＝弱上半连续.

$X,Y,Z$:通常记实 Banach 空间.

$X_+$:$X$ 中的闭凸锥;$X_-=-X_+$.

$\alpha(\cdot)$:集或映射的非紧测度.

$\chi_e$:$e$ 的特征函数.

$\Omega$:通常记 $X$ 或 $\mathrm{R}^n$ 中的开集.

$A^c$:$A$ 的补集.

$\langle u,x\rangle=u(x)$,$u\in X^*$,$x\in X$.

$\langle U,A\rangle=\{\langle u,a\rangle;u\in U,a\in A\}$.

$A\oplus B$:拓扑直和.

$[a,b]$:线段或序区间.

$|\cdot|$:$X$(或 $Y,Z$),$X^*$ 及 $L(X,Y)$中的范数.

$|\cdot|_0$:sup 范数.

$\|\cdot\|_p$：$L^p$ 范数.

$\rightharpoonup$：弱收敛.

$\overset{*}{\rightharpoonup}$：弱*收敛.

$\to(L^p)$：$L^p$ 收敛.

# 几点说明

1° 括号的用法：方括号内记述意义平行（或对偶）的条件与结论；圆括号内则是简短插入语.

2° 引证. 引证本书有关章节：§1 表示本章 §1；§1.1 表示第一章 §1；§1.1(1)表示 §1.1 中式(1). 引证文献：[1;p.1]表示文献[1]的 p.1；等等.

3° 指标用法：指标（如 $i,j,k$）变程（如 $1\leqslant i\leqslant n$）在一节中一般保持不变且只注明一次；出现于 Σ、Π、⋃、⋂ 下的指标通常省略. 给定 $x\in\mathrm{R}^n$，自动认定 $x=(x_i)$；给定 $\mathrm{R}^n$-值函数 $g$，自动认定 $g=(g_i)$.

4° 上下确界：约定 $\sup\varnothing=-\infty$，$\inf\varnothing=\infty$. 例如，若界定 $\alpha=\inf\{k;k$ 有性质 $P\}$，但有性质 $P$ 之 $k$ 不存在，则 $\alpha=\infty$.

5° 不等号的用法：$A\leqslant B\Longleftrightarrow\forall\ a\in A,b\in B:a\leqslant b$；$A<B$ 与 $A\ll B$ 仿此. 若 $u,v:D\to\mathrm{R}$，则 $u\leqslant v\Longleftrightarrow\forall\ x\in D:u(x)\leqslant v(x)$；$u<v\Longleftrightarrow\forall\ x\in D:u(x)<v(x)$.

6° 任给 $t\in[0,1]$，约定 $t'=1-t$.

7° “$x\to x_0$ 时 $f(x)$ 有性质 $P$”意味着“对充分邻近 $x_0$ 的 $x$，$f(x)$ 有性质 $P$”. 例如，“$x\to\infty$ 时 $f(x)\in A$”意指“$\exists\ a\in R,\forall\ x\geqslant a:f(x)\in A$”.

# 目　录

# 第一章　度与不动点

判定各种非线性方程的可解性，一直是非线性分析的主要课题之一．早在经典分析中，人们就发现，在一定条件下，方程的解的某种“代数个数”不因方程的连续变形而改变．这一被广泛注意到的事实引发了各种“扰动方法”的研究，这类研究最终导向更深刻的问题：能否用某种拓扑不变量来刻划方程解的“代数个数”？度理论因此而诞生．如人们所预期的，这一辉煌的理论将方程可解性的判定归结为一种拓扑不变量——度的计算．

本章给出度理论的梗概及其某些应用．

## §1　非紧测度

本章所阐述的度理论本质上依赖于一定的紧性条件，因此，必须对点集与映射的紧性作某种较精细的、“定量的”考察．

给定 $A\subset X$．如所熟知，$A$ 相对紧当且仅当 $A$ 有由直径充分小的集组成的有限覆盖．如果将“直径充分小”换成“直径充分接近于 $\alpha$”，而数 $\alpha\geqslant 0$ 需要有一定选择余地，那么就得到由数 $\alpha$ 量度的放宽了的“紧性”概念．准确地说，我们称

$$\alpha(A)\triangleq\inf\{\delta>0;\exists\ A_1,\cdots,A_n\subset X, A\subset\bigcup A_i,\mathrm{diam}A_i\leqslant\delta\}\tag{1}$$

为 $A$ 的非紧测度．实际上，$\alpha(A)$ 量度了“$A$ 对紧集的偏离”：$\alpha(A)$ 愈小，$A$ 愈接近于紧；$A$ 相对紧 $\Longleftrightarrow\alpha(A)=0$．若 $A$ 无界，则约定 $\alpha(A)=\infty$．

**1.1.1 命题**　非紧测度 $\alpha(\cdot)$ 有以下性质：(i) $\alpha(A)<\rho\Longleftrightarrow$ 存在有限分解 $A=\bigcup A_i$：$\mathrm{diam}\,A_i<\rho$．(ii) $A\subset B\Rightarrow\alpha(A)\leqslant\alpha(B)$；$\alpha(\bigcup_1^n A_i)=\max\alpha(A_i)$．(iii) $\alpha(\lambda A)=|\lambda|\alpha(A)$；$\alpha(\sum_1^n A_i)\leqslant$

$\sum\alpha(A_i)$. (iv) $\alpha(A)=\alpha(\overline{\mathrm{co}}A)$.

**证** (i)～(iii)的证明是直接的. 对于(iv)只需证$\alpha(\mathrm{co}A)\leqslant\alpha(A)$. 可设$\delta\triangleq\mathrm{diam}A<\infty$. 必定$\mathrm{diamco}A=\delta$,否则有$x,y\in\mathrm{co}A$: $|x-y|>\delta$,于是$A\not\subset\overline{B}_\delta(x)$,因而$\exists\ z\in A\backslash\overline{B}_\delta(x)$,这推出$x\in\mathrm{co}A\subset\overline{B}_\delta(z)$,与$z\bar{\in}\overline{B}_\delta(x)$矛盾. 任给$A$的覆盖$\{A_i;1\leqslant i\leqslant n\}$,今证$\alpha(\mathrm{co}A)\leqslant\max\mathrm{diam}A_i$. 可设$A_i$凸(否则以$\mathrm{co}A_i$代$A_i$),且只需证$\alpha(\mathrm{co}(\bigcup A_i))\leqslant\max\alpha(A_i)$. 不妨设$\bigcup A_i\subset B_r(0)$, $r>0$, $n=2$. $\forall\ \varepsilon>0$,取分点$0=t_0<t_1<\cdots<t_m=1$,使$t_j-t_{j-1}<\varepsilon(1\leqslant j\leqslant m)$. 约定$t'=1-t$(本书概如此),则

$$\mathrm{co}(A_1\cup A_2)\subset\bigcup_j(t_jA_1+t'_jA_2)+B_{2r\varepsilon}(0),$$

这推出$\alpha(\mathrm{co}(A_1\bigcup A_2))\leqslant\max\{\alpha(A_1),\alpha(A_2)\}+4r\varepsilon$. 令$\varepsilon\downarrow 0$得所要证. □

如所熟知,若$\{A_n\}$是$X$中非空闭集的降列且$\mathrm{diam}A_n\to0$,则$\bigcap A_n\neq\varnothing$. 下面是这一结果的推广.

**1.1.2 定理** 若$\{A_n\}$是$X$中非空闭集的降列(即$A_n\supset A_{n+1}$), $\alpha(A_n)\to0$,则$A\triangleq\bigcap A_n$是非空紧集.

**证** 因显然$\alpha(A)=0$,故只需证$A\neq\varnothing$. 任取$x_k\in A_k$,令$B_n=\{x_k;k\geqslant n\}$,则$\alpha(B_1)=\alpha(B_n)\leqslant\alpha(A_n)$,故$\alpha(B_1)=0$,因此$\{x_n\}$有收敛子列,其极限必属于$A$. □

非紧测度概念的有效应用,通常依赖于一定函数空间中点集的非紧测度的计算或估计. 对此缺少系统的方法,问题的难度因空间而异. 对于连续函数空间,已有一些较好的结果.

以下设$T$是一紧 Hausdorff 空间,在$C(T,X)$中采用 sup 范数$|\cdot|_0$(本书中未加声明时概如此). 任给$A\subset C(T,X)$,约定$A(t)=\{x(t);x\in A\}$, $A(T)=\bigcup\limits_{t\in T}A(t)$.

**1.1.3 定义** 设$A\subset C(T,X)$. 若对每个$t\in T$,关于$x\in A$一致地有$\lim\limits_{s\to t}x(s)=x(t)$(即$\forall\ \varepsilon>0$,存在$t$的邻域$V$, $\forall\ s\in V$, $x\in A$: $|x(s)-x(t)|<\varepsilon$),则说$A$等度连续. 若$\varphi:T\times D\to Y$, $D\subset X$,

$\{\varphi(\cdot,x):x\in D\}$等度连续，则说$\varphi(\cdot,x)$关于$x\in D$一致地连续.

**1.1.4 定理** 设$A\subset C(T,X)$，则$\alpha(A(T))\leqslant 2\alpha(A)$；当$A$等度连续时$t\mapsto\alpha(A(t))$连续，且

$$\alpha(A)=\max_{t\in T}\alpha(A(t))=\alpha(A(T)). \tag{2}$$

**证** $\forall\,\varepsilon>0$，取分解$A=\bigcup_1^n A_i$，使$\operatorname{diam}A_i<\alpha(A)+\varepsilon$；取$x_i\in A_i$. 由$x_i$连续与$T$紧可得分解$T=\bigcup_1^m T_j$与$t_j\in T_j$，使$|x_i(t)-x_i(t_j)|<\varepsilon(t\in T_j,1\leqslant j\leqslant m,1\leqslant i\leqslant n)$. 易验知$A(T)\subset\{x_i(t_j):1\leqslant i\leqslant n,1\leqslant j\leqslant m\}+B(0,\alpha(A)+2\varepsilon)$，由此推出$\alpha(A(T))\leqslant 2\alpha(A)+4\varepsilon$. 令$\varepsilon\downarrow 0$得$\alpha(A(T))\leqslant 2\alpha(A)$.

下面设$A$等度连续. 令$r=\sup_{x\in A}|x(t)-x(\tau)|$，则$A(\tau)\subset A(t)+\overline{B}_r(0)$. 据此易见$t\mapsto\alpha(A(t))$连续，因此存在$\mu=\max_{t\in T}\alpha(A(t))$. $\forall\,\varepsilon>0$，由$A$等度连续有分解$T=\bigcup_1^m T_j$与$t_j\in T_j$，使$|x(t)-x(t_j)|<\varepsilon(x\in A,t\in T_j,1\leqslant j\leqslant m)$. 因$\alpha(\bigcup A(t_j))\leqslant\mu$，故有$M_k\subset X(1\leqslant k\leqslant q)$，使$\bigcup A(t_j)=\bigcup M_k$，$\operatorname{diam}M_k<\mu+\varepsilon$. 任给$K=\{k_1,\cdots,k_m\}\subset\{1,\cdots,q\}$，令$A_K=\{x\in A:x(t_j)\in M_{K_j}(1\leqslant j\leqslant m)\}$，则$\operatorname{diam}A_K\leqslant\mu+3\varepsilon$，$A=\bigcup_K A_K$. 于是$\alpha(A)\leqslant\mu+3\varepsilon$，令$\varepsilon\downarrow 0$得$\alpha(A)\leqslant\mu$. 显然$\mu\leqslant\alpha(A(T))$. 设$A_i$依上段，则$A(T)\subset\bigcup_{i,j}A_i(t_j)+B_\varepsilon(0)$，这推出

$$\alpha(A(T))\leqslant\max_{i,j}\operatorname{diam}A_i(t_j)+2\varepsilon\leqslant\alpha(A)+3\varepsilon.$$

令$\varepsilon\downarrow 0$得$\alpha(A(T))\leqslant\alpha(A)$，于是(2)式得证. □

由1.1.4直接推出：若$A$等度连续，则$A$相对紧$\Longleftrightarrow\forall\,t\in T$：$A(t)$相对紧. 因此可以说，1.1.4是著名的Arzela-Ascoli定理[75;1.3.5]的推广.

将1.1.4用于$A=\{\varphi(\cdot,x):x\in D\}$得出：

**1.1.5 推论** 设$D\subset X$，$\varphi:T\times D\to Y$，$\varphi(\cdot,x)$关于$x\in D$一致地连续，则$\alpha(\varphi(T\times D))=\max_{t\in T}\alpha(\varphi(t,D))$.

以下设$(T,\mu)$是一测度空间. 在Lebesgue空间$L^p(T,X)$中解

决类似于 1.1.4 的问题是更困难的. 下面给出几个较简单的结果. 记 $S_A = \left\{\int_T x\mathrm{d}\mu : x \in A\right\}$.

**1.1.6命题** 设 $\mu T<\infty$, $A\subset L^1(T,X)$ 有界, 则 $\alpha(S_A)\leqslant \mu T\cdot\alpha(A(T))$.

**证** 由 1.1.1(iv), 只要证 $S_A\subset\mu T\,\overline{\mathrm{co}}A(T)$. 可设 $\mu T>0$. 取定 $x\in A$, 今证 $(\mu T)^{-1}\int_T x\mathrm{d}\mu\in\overline{\mathrm{co}}A(T)$. $\forall\ \varepsilon>0$, 取可数值函数 $y=\sum a_i\chi_{e_i}$, 使 $\mu(T\backslash T_0)=0$, $T_0=\bigcup e_i$, $|x(t)-y(t)|<\varepsilon(\forall\ t\in T_0)$. 取 $\delta\in(0,\varepsilon(\mu T)^2/2)$, 使当 $S\subset T$, $\mu S<\delta$ 时 $\int_S|y|\mathrm{d}\mu<\varepsilon\mu T$. 令 $Q=\bigcup_1^n e_i$, $z=\sum_1^n a_i\chi_{e_i}$, 设 $\mu(T\backslash Q)<\min\{\delta,\mu T/2\}$, 则

$$\left|\int_T z\mathrm{d}\mu-\int_T y\mathrm{d}\mu\right|=\int_{T\backslash Q}|y|\mathrm{d}\mu<\varepsilon\mu T.$$

取 $t_i\in e_i$, 则 $(\mu Q)^{-1}\sum_1^n x(t_i)\mu e_i\in\mathrm{co}A(T)$, 而

$$\begin{aligned}
&\left|\frac{1}{\mu T}\int_T x\mathrm{d}\mu-\frac{1}{\mu Q}\sum_1^n x(t_i)\mu e_i\right|\\
&\leqslant\frac{1}{\mu T}\int_T|x-y|\mathrm{d}\mu+\frac{1}{\mu T}\left|\int_T y\mathrm{d}\mu-\int_T z\mathrm{d}\mu\right|\\
&\quad+\left(\frac{1}{\mu T}-\frac{1}{\mu Q}\right)\int_T|z|\mathrm{d}\mu+\frac{1}{\mu Q}\left|\int_T z\mathrm{d}\mu-\sum_1^n x(t_i)\mu e_i\right|\\
&<\varepsilon+\varepsilon+\frac{2\delta}{(\mu T)^2}(|x|_1+\varepsilon\mu T)+\varepsilon\\
&<\varepsilon(3+|x|_1+\varepsilon\mu T),
\end{aligned}$$

由此得出 $(\mu T)^{-1}\int_T x\mathrm{d}\mu\in\overline{\mathrm{co}}A(T)$. □

任给 $A\subset X$, 称

$$\beta(A)\triangleq\inf\{\delta>0:\exists\ x_1,x_2,\cdots,x_n\in X, A\subset\bigcup B_\delta(x_i)\}\quad(3)$$

为 $A$ 的"球非紧测度". $\beta$ 具有类似于 $\alpha$ 的性质, 且

$$\beta(A)\leqslant\alpha(A)\leqslant2\beta(A).\qquad(4)$$

**1.1.7引理** 设 $X_n=\mathrm{span}\{a_1,a_2,\cdots,a_n\}$, $X=\overline{\{a_n:n\geqslant1\}}$, $A=\{x_n:n\geqslant1\}\subset X$, 则 $\beta(A)=\lim\limits_n\overline{\lim\limits_m}d(x_m,X_n)$.

证明是直接的.

**1.1.8 定理**(Heinz, 1983[68]) 设 $A=\{x_n: n\geqslant 1\}\subset L^1(T, X)$ 一致可积，即∃ $g\in L^1(T)$, $\forall\ n\geqslant 1$: $|x_n(t)|\leqslant g(t)$，则

$$\alpha(S_A)\leqslant 2\int_T \beta(A(t))\mathrm{d}t, \mathrm{d}t \text{ 记 } \mathrm{d}\mu(t). \tag{5}$$

**证** 因每个 $x_n$ 几乎可分值[75;3.2.2]，不妨设 $X$ 可分. 令 $X=\overline{\{a_n: n\geqslant 1\}}$, $X_n=\mathrm{span}\{a_1,a_2,\cdots,a_n\}$，用 1.1.7 及 Levi 定理与 Fatou 定理得

$$\begin{aligned}\alpha(S_A)\leqslant 2\beta(S_A) &= 2\lim_n \overline{\lim_m} d\left(\int_T x_m \mathrm{d}t, X_n\right)\\ &\leqslant 2\lim_n \overline{\lim_m}\int_T d(x_m(t), X_n)\mathrm{d}t\\ &\leqslant 2\int_T \lim_n \overline{\lim_m} d(x_m(t), X_n)\mathrm{d}t\\ &= 2\int_T \beta(A(t))\mathrm{d}t.\end{aligned}$$

□

## §2 集压缩映射

给定 $F\in C(D,Y)$, $D\subset X$. 熟知

$$\mathrm{Lip}F \triangleq \inf\{k>0: |Fx-Fy|\leqslant k|x-y| (\forall\ x,y\in D)\} \tag{1}$$

刻划了 $F$ 的"度量压缩性". $F$ 是 Lipschitz 映射(本书中简写作"$F$ 为 Lip")$\Longleftrightarrow \mathrm{Lip}F<\infty$; $F$ 是非扩张映射 $\Longleftrightarrow \mathrm{Lip}F\leqslant 1$; $F$ 是压缩映射 $\Longleftrightarrow \mathrm{Lip}F<1$. 关于压缩映射的以下不动点定理是熟知的:

**1.2.1 定理**(Banach) 若 $F: D\to D$, $D\subset X$ 闭, $\mathrm{Lip}F<1$，则 $F$ 有唯一不动点 $x\in D$，即 $Fx=x$.

鉴于 1.2.1 在现代分析中的巨大价值，出现了放宽条件以扩充其应用的种种尝试. 最富有启发性的推广是以所谓集压缩映射代替压缩映射. 仿照(1)，定义一个刻划 $F$ 的"集压缩性"的量:

$$\alpha(F)=\inf\{k>0: \alpha(FA)\leqslant k\alpha(A) (\forall\ A\subset D)\}, \tag{2}$$

称 $\alpha(F)$ 为 $F$ 的"非紧测度". 与前述的"Lipschitz 概念"相对应，约

定 $F$ 是 $\alpha$-Lipschitz 映射$\Longleftrightarrow\alpha(F)<\infty$;$F$ 是集压缩映射$\Longleftrightarrow\alpha(F)<1$. 若 $\alpha(F)=0$($\Longleftrightarrow F$ 映有界集为相对紧集),则称 $F$ 为全连续映射;若 $FD$ 相对紧,则称 $F$ 为紧映射. 分别以 SC 与 $\mathscr{K}$ 记集压缩映射类与紧映射类;SC($D$)记从$D$到$X$的集压缩映射之全体,$\mathscr{K}(D)$仿此.

若 $F$ 映有界集为有界集,则称 $F$ 为有界映射. 显然 $\alpha(F)<\infty\Rightarrow F$ 为有界映射. 若 $\dim Y<\infty$,则 $F\in C(D,Y)$全连续$\Longleftrightarrow F$ 有界;若 $\dim X<\infty$且 $D$ 闭,则 $F:D\to Y$ 全连续$\Longleftrightarrow F$ 连续. 可见全连续(更不必说集压缩)映射概念本质上是为无限维空间中的映射而设的.

$\alpha(\cdot)$的以下性质虽然简单但很有用.

**1.2.2命题** (i)$\alpha(FA)\leqslant\alpha(F)\alpha(A)$($\forall\ A\subset D$). (ii)半范性:$\alpha(F+G)\leqslant\alpha(F)+\alpha(G)$,$\alpha(\lambda F)=|\lambda|\alpha(F)$. (iii)次乘法性:$\alpha(G\circ F)\leqslant\alpha(G)\alpha(F)$. 以上假定 $F+G$,$G\circ F$ 有意义. (iv)$\alpha(F)\leqslant\mathrm{Lip}F$.

对于 $T\in L(X,Y)$,由 1.2.2(iv)有 $\alpha(T)\leqslant\mathrm{Lip}T=|T|$;$T$ 是紧线性算子$\Longleftrightarrow\alpha(T)=0$. 约定 $CL(X)=\{T\in L(X):\alpha(T)=0\}$.

**1.2.3 命题** 若 $x\in D^\circ$,$T=F'(x)$,或 $T=F'(\infty)$(这意味着 $T\in L(X,Y)$,当$|x|\to\infty$时 $Fx$ 有定义且$|Fx-Tx|=o(|x|)$),则 $\alpha(T)\leqslant\alpha(F)$. 特别,当 $F$ 全连续时 $T$ 是紧线性算子.

**证** 首先设 $T=F'(x)$,$x\in D^\circ$. $\forall\ \varepsilon>0$,取 $\delta>0$,使当$|h|\leqslant\delta$时$|\Delta F(x,h)-Th|\leqslant\varepsilon|h|$. 任给有界集$A\subset X$,取$r>0$:$A\subset B_r(0)$,则 $r^{-1}\delta A\subset B_\delta(0)$,于是

$$T\left(\frac{\delta}{r}A\right)\subset F\left(x+\frac{\delta}{r}A\right)-Fx+B_{\varepsilon\delta}(0);$$

$$\alpha(TA)\leqslant\frac{r}{\delta}\left[\alpha\left(F\left(x+\frac{\delta}{r}A\right)\right)+2\varepsilon\delta\right]$$

$$\leqslant\frac{r}{\delta}\alpha(F)\alpha\left(x+\frac{\delta}{r}A\right)+2r\varepsilon=\alpha(F)\alpha(A)+2r\varepsilon.$$

令 $\varepsilon\downarrow 0$ 得 $\alpha(TA)\leqslant\alpha(F)\alpha(A)$,因此 $\alpha(T)\leqslant\alpha(F)$.

其次设 $T=F'(\infty)$. $\forall\ \varepsilon>0$,取 $\rho>0$,使当$|x|\geqslant\rho$ 时$|Fx-Tx|$

$\leqslant\varepsilon|x|$. 仍设 $A\subset B_r(0)$，令 $A'=A\cap B_c(0)$，$A''=A\setminus A'$. 显然 $\alpha(TA')\leqslant 2|T|\varepsilon$. 由"$x\in\rho\varepsilon^{-1}A''\Rightarrow|x|\geqslant\rho$"得 $T(\rho\varepsilon^{-1}A'')\subset F(\rho\varepsilon^{-1}A'')+B_{\rho r}(0)$，于是

$$\begin{aligned}\alpha(TA'')&\leqslant\frac{\varepsilon}{\rho}\left[\alpha\left(F\left(\frac{\rho}{\varepsilon}A''\right)\right)+2\rho r\right]\\&\leqslant\alpha(F)\alpha(A)+2r\varepsilon.\end{aligned}$$

因此 $\alpha(TA)\leqslant\alpha(F)\alpha(A)+2\varepsilon(r+|T|)$，这同样得出所要证. □

**1.2.4 命题** 设 $F=I-f:D\to X$ 是集压缩的，$D$ 是有界闭集，$K\subset X$ 紧，$A\subset D$ 闭，则 $f^{-1}K$ 紧而 $fA$ 闭.

**证** 若 $C\triangleq f^{-1}K$ 非紧，则 $\alpha(C)>0$，这与 $C\subset FC+K$ 一起推出 $\alpha(C)\leqslant\alpha(FC)<\alpha(C)$，得出矛盾. 若 $x_n\in A$，$f(x_n)\to y$，则 $M\triangleq\{f(x_n):n\geqslant 1\}\cup\{y\}$ 紧，从而 $f^{-1}M$ 紧，于是有 $x_{n_i}\to x\in A$，因此 $y=f(x)\in fA$. 这证得 $fA$ 闭. □

**1.2.5 扩张定理** 设 $F\in C(D,Y)$，$D\subset X$ 闭. (i) $F$ 有扩张 $\widetilde{F}\in C(X,Y)$，使得 $\widetilde{F}X\subset\mathrm{co}FD$；当 $F\in\mathscr{K}$ 时亦必 $\widetilde{F}\in\mathscr{K}$. (ii) 若 $F$ 全连续，则 $F$ 有全连续的扩张 $\widetilde{F}:X\to Y$，使得 $\widetilde{F}X\subset\mathrm{co}FD$. (iii) 若 $D=\overline{B}_r(x_0)$，或 $D$ 闭凸而 $X$ 为 Hilbert 空间，则 $F$ 有扩张 $\widetilde{F}\in C(X,Y)$，使得 $\widetilde{F}X=FD$，且 $\alpha(\widetilde{F})=\alpha(F)$.

**证** (i) 令 $r_z=2^{-1}d(z,D)$，则 $\{B(z,r_z):z\in D^c\}$ 是 $D^c$ 的开覆盖，于是有 $D^c$ 的局部有限开覆盖 $\{U_\alpha\}$，使得 $U_\alpha\subset B(z_\alpha,r_\alpha)$，$z_\alpha\in D^c$，$r_\alpha=r_{z_\alpha}$（这用到 $D^c$ 的仿紧性，参考[97;Ch. 4]）. 取 $x_\alpha\in D\cap B(z_\alpha,3r_\alpha)$，令 $\varphi_\alpha(x)=d(x,U_\alpha^c)$. 定义

$$\widetilde{F}x=\sum_\alpha\varphi_\alpha(x)Fx_\alpha\Big/\sum_\alpha\varphi_\alpha(x),\quad x\in D^c;\tag{3}$$

而置 $\widetilde{F}|D=F$，则 $\widetilde{F}:X\to Y$ 是 $F$ 的扩张，且 $\widetilde{F}X\subset\mathrm{co}FD$，$\widetilde{F}$ 在 $D^c$ 内连续. 今证 $\widetilde{F}$ 在任一点 $a\in D$ 连续. $\forall\ \varepsilon>0$，取 $\delta>0$：$F(D\cap B_{5\delta}(a))\subset B_\varepsilon(Fa)$. $\forall\ x\in D^c\cap B_\delta(a)$，设 $x\in U_\alpha$，则

$$|x_\alpha-a|\leqslant|x_\alpha-z_\alpha|+|z_\alpha-x|+\delta<4r_\alpha+\delta<5\delta,$$

从而 $Fx_\alpha\in B_\varepsilon(Fa)$. 于是依(3)有 $\widetilde{F}x\in B_\varepsilon(Fa)$，因此 $\widetilde{F}B_\delta(a)\subset$

$B_\varepsilon(Fa)$.

(ii)令 $B_n=\overline{B}_n(0)$，$D_n=D\cup B_n$，不妨设 $D\cap B_1\neq\varnothing$. 由已证之(i)，$F|D\cap B_1$ 有扩张 $\widetilde{F}_1\in C(X,Y)$：$\widetilde{F}_1X\subset \mathrm{co}F(D\cap B_1)$. 依 $F_1|D=F$，$F_1|B_1=\widetilde{F}_1|B_1$ 定义出 $F$ 的扩张 $F_1:D_1\to Y$，$F_1D_1\subset\mathrm{co}FD$，$F_1$ 全连续. 同理，$F_1$ 有全连续扩张 $F_2:D_2\to Y$，$F_2D_2\subset\mathrm{co}F_1D_1\subset\mathrm{co}FD$，…，如此得一序列 $\{F_n\}$. 依 $\widetilde{F}|B_n=F_n|B_n(n=1,2,\cdots)$定义的 $\widetilde{F}:X\to Y$ 即合所求.

(iii)若存在从 $X$ 到 $D$ 的收缩 $P$(即 $P\in C(X,D)$，$P|D=\mathrm{id}$)，使得 $\alpha(P)\leqslant1$，则 $\widetilde{F}=F\cdot P$ 显然合于要求. 对 $D=\overline{B}_r(x_0)$，取 $P$ 为径向收缩，即当 $x\in D'$ 时令 $Px=x_0+r(x-x_0)/|x-x_0|$，而置 $P|D=\mathrm{id}$. 因对任给 $A\subset X$ 有 $PA\subset\mathrm{co}(A\cup\{x_0\})$，故 $\alpha(PA)\leqslant\alpha(A)$，因此 $\alpha(P)\leqslant1$. 若 $X$ 是 Hilbert 空间而 $D$ 闭凸，则 5.11.2 将指明有收缩 $P:X\to D$ 满足 $\mathrm{Lip}P\leqslant1$. □

**1.2.6 逼近定理** 设 $D\subset X$ 有界，$F:D\to Y$ 为紧映射. 则 $\forall\,\varepsilon>0$，$\exists\,G\in C(D,Y)$：$\dim\mathrm{span}G(D)<\infty$ 且 $|F-G|_0\leqslant\varepsilon$(约定 $|F|_0=\sup\limits_x|Fx|$，本书概如此).

**证** 由 $FD$ 相对紧有有限集 $\{y_i\}\subset Y$，使 $FD\subset\bigcup B_\varepsilon(y_i)$. 令 $\varphi_i(x)=d(Fx,Y\backslash B_\varepsilon(y_i))$，定义

$$Gx=\sum\varphi_i(x)y_i/\sum\varphi_i(x),\ x\in D \tag{4}$$

(可与(3)对照). 显然 $G\in C(D,Y)$，$GD\subset\mathrm{span}\{y_i\}$. 令 $\lambda_i=\varphi_i/\sum\limits_j\varphi_j$，则 $\sum\lambda_i=1$；对任给 $x\in D$ 有

$$|Fx-Gx|\leqslant\sum\lambda_i(x)|Fx-y_i|<\varepsilon. \qquad \square$$

约定 $J=[0,1]$. 称任何 $H\in C(J\times D,Y)$为同伦；若 $H$ 是紧映射，则称 $H$ 为紧同伦；若有 $k<1$，使对任给 $A\subset D$ 有 $\alpha(H(J\times A))\leqslant k\alpha(A)$，则称 $H$ 为 SC-同伦. 对于一个同伦 $H(t,x)$，通常以 $H_t$ 记 $H(t,\cdot)$，并说 $H_t$ 是从 $H_0$ 到 $H_1$ 的同伦. 同伦的意义在于，若某个与映射有关的量 $q$“同伦不变”，则可利用等式 $q(H_0)=q(H_1)$

来简化 $q(H_0)$ 或 $q(H_1)$ 的计算. 这一思想在本章有本质意义.

**1.2.7 命题** 设 $D\subset X$, $H: J\times D\to Y$. (i) 若 $H(t,\cdot)$ 连续, $H(\cdot,x)$ 关于 $x\in D$ 一致地连续, $\alpha(H_t)\leqslant k<1$, 则 $H$ 是 SC-同伦. (ii) 若 $H$ 是 SC-同伦, $h(t,x)=x-H(t,x)$, $S\subset D$ 是有界闭集, 则 $h(J\times S)$ 闭.

**证** (i) 显然 $H$ 连续, 而由 1.1.5 有 $\alpha(H(J\times A))\leqslant \sup_t \alpha(H_t(A))\leqslant k\alpha(A)$ $(A\subset D)$, 因此 $H$ 是 SC-同伦.

(ii) 设 $y_n\triangleq h(t_n,x_n)\to y$, $(t_n,x_n)\in J\times S$, 可设 $t_n\to t\in J$. 令 $A=\{x_n\}$, $B=\{y_n\}$, 则 $A\subset B+H(J\times A)$, $\alpha(B)=0$. 若 $\alpha(A)>0$, 则 $\alpha(A)\leqslant\alpha(H(J\times A))<\alpha(A)$, 得出矛盾. 因此 $\alpha(A)=0$, 故不妨设 $x_n\to x\in S$, 于是 $y=h(t,x)\in h(J\times S)$, $h(J\times S)$ 是闭的. □

**1.2.8 推论** 设 $D\subset X$ 有界, $F,G: D\to Y$ 是集压缩映射, 则 $H_t\triangleq t'F+tG$ $(t\in J)$ 是从 $F$ 到 $G$ 的 SC-同伦.

**注** 对于同伦限定参数 $t\in J$ 并无本质意义, 今后将依需要自由地以其它紧区间取代 $J$.

## §3 拓扑度

现在进入本章主要课题——度理论的讨论. 简要说来, 度理论的目标是, 构成一个称为度的整值函数 $d(\cdot,\cdot,\cdot)$, 使对任给有界开集 $\Omega\subset X$, $F=I-f\in SC(\overline{\Omega})$, $y\in X\setminus f(\partial\Omega)$, $d(f,\Omega,y)$ 有定义且表出方程 $f(x)=y$ 在 $\Omega$ 内的解的某种"代数个数". 这一问题经历了长久的探索之后终获解决, 其详细解答颇不简单, 但其最终结论倒极明了且令人深感满意. 鉴于对度的构成已有极完善的表述 (参看[34,43,55,202]), 本书避繁就简, 以直接陈述度的基本性质作为出发点, 演绎地展开度的基本理论, 以期尽快达到较深入的可用结论. 这一选择不免留下一些缺憾, 但似乎足以使主要关心度的应用的读者满足. 因此, 让我们直接面对不加证明地使用的以下结论:

**1.3.1 定理** 存在唯一整值函数 $d(\cdot,\cdot,\cdot)$，使对任给有界开集 $\Omega\subset X$，$F=I-f\in \mathrm{SC}(\overline{\Omega})$，$y\in X\setminus f(\partial\Omega)$，$d(f,\Omega,y)$ 有定义，且有性质：$(\mathrm{D}_1)$ 单位性：$d(I,\Omega,y)=1(\forall\ y\in\Omega)$；$(\mathrm{D}_2)$ 可加性：若 $\Omega_i(1\leqslant i\leqslant n)$ 是 $\Omega$ 的互不相交开子集，$y\in X\setminus f(\overline{\Omega}\setminus\bigcup\Omega_i)$，则 $d(f,\Omega,y)=\sum_i d(f,\Omega_i,y)$；$(\mathrm{D}_3)$ 同伦不变性：若 $H:J\times\overline{\Omega}\to X$ 是 SC-同伦，$y\in C(J,X)$，$y(t)\overline{\in}(I-H_t)(\partial\Omega)(\forall\ t\in J)$，则 $d(I-H_t,\Omega,y(t))$ 与 $t$ 无关.

称 1.3.1 中的函数 $d$(亦写作 deg)为度或拓扑度. 若限定 $F\in\mathscr{K}$ 或 $X=\mathrm{R}^n$，则相应的度分别称为 Leray-Schauder 度(简称 LS 度)与 Brouwer 度.

$(\mathrm{D}_3)$ 是最重要的度性质，它表明 $d(f,\Omega,y)$ 不因 $f,y$ 的"连续变形"而改变. $(\mathrm{D}_2)$ 则表明可通过适当"分割"$\Omega$ 来计算 $d(f,\Omega,y)$. 利用 $(\mathrm{D}_1)\sim(\mathrm{D}_3)$ 适当地改变 $f,\Omega,y$，有可能简化度的计算.

由 $(\mathrm{D}_1)\sim(\mathrm{D}_3)$ 可推得度的一些更进一步的性质.

**1.3.2 推论** 度 $d(\cdot,\cdot,\cdot)$ 有以下性质：$(\mathrm{D}_4)$ 切除性：若 $A\subset\overline{\Omega}$ 闭，$y\in X\setminus f(A\cup\partial\Omega)$，则 $d(f,\Omega,y)=d(f,\Omega\setminus A,y)$；$d(f,\Omega,y)\neq 0\Rightarrow f^{-1}(y)\neq\varnothing$. $(\mathrm{D}_5)$ 若 $y\in X\setminus[f(x),g(x)](\forall\ x\in\partial\Omega)$(特别，若 $|Fx-Gx|<|y-f(x)|(\forall\ x\in\partial\Omega)$ 或 $F|\partial\Omega=G|\partial\Omega$，$y\in X\setminus f(\partial\Omega)$)，则 $d(f,\Omega,y)=d(g,\Omega,y)$. $(\mathrm{D}_6)$ 归约性质：若 $F\overline{\Omega}\subset Y\subset X$，$Y$ 是闭子空间，$y\in Y\setminus f(\partial\Omega)$，则 $d(f,\Omega,y)=d(f,Y\cap\Omega,y)$(右端的 $f$ 理解为 $f|Y\cap\overline{\Omega}$). $(\mathrm{D}_7)$ $d(f,\Omega,\cdot)$ 在 $X\setminus f(\partial\Omega)$ 的各分支内取常值，在无界分支内为零. 以上 $F=I-f\in\mathrm{SC}(\overline{\Omega})$，$G=I-g\in\mathrm{SC}(\overline{\Omega})$.

**证** 直接用 $(\mathrm{D}_2)$ 推出 $(\mathrm{D}_4)$. 取 $H_t=t'F+tG$(参考 1.2.8)从 $(\mathrm{D}_3)$ 推出 $(\mathrm{D}_5)$. 因 $f(\partial\Omega)$ 闭(1.2.4)，故 $V\triangleq X\setminus f(\partial\Omega)$ 开，从而 $V$ 的任一分支 $P$ 为路连通开集[75;1.4.3]，据此易从 $(\mathrm{D}_3)$ 推出 $d(f,\Omega,\cdot)$ 在 $P$ 内为常数. 若 $P$ 无界，则必 $\exists\ y\in P\setminus f\overline{\Omega}$(注意 $f\overline{\Omega}$ 有界!)，于是由 $(\mathrm{D}_4)$ 有 $d(f,\Omega,y)=0$. 这证得 $(\mathrm{D}_7)$. 至于 $(\mathrm{D}_6)$，尽管

它确为$(D_1)\sim(D_3)$的逻辑推论，但其证明有赖于LS度与Brouwer度的某种程度的展开，只好从略. □

$(D_4)$表明，若$d(f,\Omega,y)\neq 0$，则方程$f(x)=y$在$\Omega$内有解，度的应用价值主要基于此. $(D_5)(D_7)$表明$d(f,\Omega,y)$连续地依赖于$f,y$，通常利用这一点来简化度的计算. 例如，用$(D_5)$可得如下结果：

**1.3.3 命题** 设$F\in SC(\Omega)$，$x_0\in\Omega$，$Fx_0=x_0$，$A=F'(x_0)$存在且不以1为特征值，则对适当小的$r>0$有

$$d(I-F,B_r(x_0),0)=d(I-A,B_r(0),0). \qquad (1)$$

**证** 不妨设$x_0=0$. 因$A\in SC(1.2.3)$，故$(I-A)(\partial B_1(0))$闭(1.2.4)，这结合$A$无特征值1得出$2\varepsilon\triangleq\inf\limits_{|x|=1}|x-Ax|>0$. 取$r>0$充分小，使当$|x|\leqslant r$时$|Fx-Ax|\leqslant\varepsilon|x|$. 于是$\forall\ x\in\partial B_r(0)$：$|Fx-Ax|\leqslant\varepsilon|x|=\varepsilon r<2\varepsilon r\leqslant|x-Ax|$，因此由$(D_5)$得出(1). □

在本章及下章中引用的$(D_i)(1\leqslant i\leqslant 7)$总是指1.3.1与1.3.2中所述的度性质而不另作说明.

应用度理论于方程问题的一种简单方式是：指明$d(f,\Omega,y)=\sum d(f,\Omega_i,y)(i=1,2)$且$d(f,\Omega,y)$与$d(f,\Omega_1,y)$分别为1,0(或0,1)，由此得出$d(f,\Omega_2,y)\neq 0$，于是方程$f(x)=y$在$\Omega_2$中有解. 因此判定$d(f,\Omega,y)=1$或0有其重要性，以下两个定理正为此而设.

**1.3.4 定理** 设$x_0\in\Omega$，$\overline{\Omega}\subset D\subset X$，$F=I-f\in SC(D)$. 以下每个条件蕴涵$d(f,\Omega,0)=1$：

$(H_1)$ Leray-Schauder条件(LS)：$(\lambda,x)\in J\times\partial\Omega\Rightarrow x-x_0\neq\lambda(Fx-x_0)$；

$(H_2)$ $\Omega=B_r(x_0)$，$r>0$充分小，$Fx_0=x_0$，$A=F'(x_0)$存在且无$\geqslant 1$的实特征值；

$(H_3)$ $D=X$，$\Omega=B_\rho(0)$，$\rho>0$充分大，$B=F'(\infty)$存在且无$\geqslant 1$的实特征值；

($H_4$) $\forall\ x\in\partial\Omega$: $|Fx-x_0|\leqslant|x-x_0|$ 且 $Fx\neq x$.

**证** 可设 $x_0=0$. 当条件($H_1$)满足时由($D_3$)($D_1$)有

$$d(f,\Omega,0)=d(I-tF,\Omega,0)=d(I,\Omega,0)=1\ (t\in J).$$

若条件($H_2$)满足，则由 1.3.3 有 $d(f,\Omega,0)=d(I-A,\Omega,0)$. $A$ 无 $\geqslant1$ 的特征值$\Rightarrow A$ 满足(LS)(取 $x_0=0$)，因此 $d(I-A,\Omega,0)=1$. 类似于 1.3.3 可证($H_3$)$\Rightarrow d(f,\Omega,0)=d(I-B,\Omega,0)=1$. 最后，显然($H_4$)$\Rightarrow$($H_1$). □

**1.3.5定理** 设 $F=I-f\in SC(\overline{\Omega})$，则以下每个条件蕴涵 $d(f,\Omega,0)=0$:

($H'_1$) $\exists\ e\in X\backslash\{0\}$: $f(\partial\Omega)\cap\mathbf{R}_+e=\varnothing$;

($H'_2$) $f(\partial\Omega)$含于 $X$ 的某真子空间，$0\overline{\in}f(\partial\Omega)$;

($H'_3$) $\dim X=\infty$, $F\in\mathscr{K}$; $d(0,F(\partial\Omega))>0$, $\forall\ \lambda\geqslant1$: $\partial\Omega\cap\mathrm{Fix}(\lambda F)=\varnothing$;

($H'_4$) $\dim X=\infty$, $F\in\mathscr{K}$, $0\overline{\in}\partial\Omega$, $\forall\ x\in\partial\Omega$: $|x|\leqslant|Fx|$ 且 $Fx\neq x$.

**证** 令 $F_t=F+te$，则条件($H'_1$)推出 $\partial\Omega\cap\mathrm{Fix}(F_t)=\varnothing$ ($\forall\ t\geqslant0$)；而当 $t$ 充分大时 $\overline{\Omega}\cap\mathrm{Fix}F_t=\varnothing$，于是由($D_3$)($D_4$)有 $d(f,\Omega,0)=d(I-F_t,\Omega,0)=0$. 其次显然($H'_2$)$\Rightarrow$($H'_1$).

若条件($H'_3$)满足，则 $\varepsilon\triangleq\inf\{|tx-Fx|:(t,x)\in J\times\partial\Omega\}>0$ (否则有$(t_n,x_n)\in J\times\partial\Omega$: $t_nx_n-Fx_n\to0$. 可设 $t_n\to t\in J$, $Fx_n\to y$. 由 $d(0,F(\partial\Omega))>0$ 推出 $y\neq0$，于是 $t>0$, $x_n\to x=t^{-1}y\in\partial\Omega$, $tx=Fx$，与 $\partial\Omega\cap\mathrm{Fix}(t^{-1}F)=\varnothing$矛盾). 由 1.2.6，有有限维空子空间 $Y\subset X$ 与 $G\in C(\overline{\Omega},Y)$: $|F-G|_0<\varepsilon$. 取 $e\in X\backslash Y$ ($\dim X=\infty$用于此!)，令 $Z=\mathrm{span}(Y\cup\{e\})$, $h(t,x)=tx-Gx$, $(t,x)\in J\times\overline{\Omega}$，则当$(t,x)\in J\times\partial\Omega$ 时 $|h(t,x)|\geqslant|tx-Fx|-|F-G|_0>0$，于是由($D_5$)($D_6$)($D_3$)有

$$d(f,\Omega,0)=d(I-G,\Omega,0)=d(I-G,Z\cap\Omega,0)$$
$$=d(h_t,Z\cap\Omega,0)=d(-G,Z\cap\Omega,0).$$

用条件$(H'_2)$得出$d(-G,Z\cap\Omega,0)=0$. 最后，显然$(H'_4)\Rightarrow(H'_3)$.

□

## §4 不动点定理

无论在理论上或在方法上，将方程问题转化为不动点问题是可取的，而且已成为一种通行的模式. 不动点定理作为一类基本的分析工具已确立其牢固地位.

约定$\mathrm{Fix}F$记映射$F$的不动点集，即$\mathrm{Fix}F=\{x:Fx=x\}$. 显然$\mathrm{Fix}F=(I-F)^{-1}(0)$，因此若$d(I-F,\Omega,0)\neq0$，则$\Omega\cap\mathrm{Fix}F\neq\varnothing$. 这一简单事实乃是由度理论导出不动点定理的基本依据. 从方法上考虑，在形成实用的不动点定理时，必须提炼出某些便于验证的条件，而避开度的直接计算. 一个基于1.3.4的基本结果是：

**1.4.1 定理**(Leray-Schauder)　设$\Omega\subset X$为有界开集，$0\in\Omega$，$F\in SC(\overline{\Omega})$. 若$(\lambda,x)\in(0,1)\times\partial\Omega\Rightarrow x\neq\lambda Fx$，则$\overline{\Omega}\cap\mathrm{Fix}F\neq\varnothing$.

1.4.1中的条件“$(\lambda,x)\in(0,1)\times\partial\Omega\Rightarrow x\neq\lambda Fx$”可代以如下条件之一：

(i) Rothe 条件：$|Fx|\leqslant|x|$ ($\forall\ x\in\partial\Omega$，下同)；

(ii) Altman 条件：$|Fx|^2\leqslant|x|^2+|x-Fx|^2$；

(iii) Krasnoselskii 条件：$X$是Hilbert空间，$(Fx,x)\leqslant|x|^2$.

1.4.1是形式上很简单而具足够强度的不动点定理之一. 不过，其条件之验证并不总是容易的. 有时我们希望使用形式上更简单或直观上更明确的结果. 下面的定理正合此要求.

**1.4.2 定理**　设$D\subset X$是有界闭凸集，$F\in SC(D)$，$F(\partial D)\subset D$，则$\mathrm{Fix}F\neq\varnothing$.

**证**　首先设$\Omega\triangleq D^\circ\neq\varnothing$，不妨设$0\in\Omega$. $F$必满足1.4.1之条件. 否则有$\lambda\in(0,1)$，$x\in\partial\Omega\subset\partial D$：$\lambda^{-1}x=Fx\in D$. 若$y$充分邻近$x$，则$z\triangleq(y-x)/\lambda'\in D$，于是$y=\lambda Fx+\lambda' z\in D$，因而$x\in D^\circ$，与$x\in\partial D$矛盾. 于是$\mathrm{Fix}F\neq\varnothing$.

其次设 $D^\circ=\varnothing$，于是 $FD\subset D$. 若 $F\in\mathscr{K}$，则由 1.2.5 有 $F$ 的紧扩张 $\tilde{F}:X\to X$，$\tilde{F}X\subset \text{co}FD\subset D$. 取 $\overline{B}_r(0)\supset D$，对 $\tilde{F}|\overline{B}_r(0)$ 应用上段结论得 $\text{Fix}F=\text{Fix}\tilde{F}\neq\varnothing$. 若 $F\bar{\in}\mathscr{K}$，取 $x_0\in D$，令 $\beta=\{A:x_0\in A\subset D, A$ 闭凸，$FA\subset A\}$，则 $D\in\beta$；$B\triangleq\bigcap_{A\in\beta}A\in\beta$；$C\triangleq\overline{\text{co}}(\{x_0\}\cup FB)\in\beta$. 易见 $B=C$，因此 $\alpha(B)=\alpha(FB)\leqslant\alpha(F)\alpha(B)$，这推出 $\alpha(B)=0$，从而 $B$ 紧. 对 $F:B\to B$ 应用已证结论得 $\text{Fix}F\neq\varnothing$. □

1.4.2 涵盖了好几个著名的不动点定理.

**1.4.3 推论** 设 $D\subset X$ 为有界闭凸集、$F\in C(D,D)$，则以下每个条件蕴涵 $\text{Fix}F\neq\varnothing$：(i) $\alpha(F)<1$；(ii) $F\in\mathscr{K}$；(iii) $D$ 紧；(iv) $X=\mathbb{R}^n$.

**注** 1.4.3 中条件(i)(ii)(iv)分别对应 Darbo 定理、Schauder 定理与 Brouwer 定理，它们是最常用的不动点定理之一. 其中 Darbo 定理可称之为“集压缩映射原理”，它是“压缩映射原理”(1.2.1)的优美类似.

称拓扑空间 $T$ 有不动点性质(FPP)，若 $\forall F\in C(T,T)$：$\text{Fix}F\neq\varnothing$. 1.4.3 表明 Banach 空间中的紧凸集有 FPP；有界闭凸集却未必如此.

**1.4.4 例** 无限维 Hilbert 空间 $X$ 中的闭单位球 $B$ 没有 FPP. 可设 $X=l^2$，定义

$$F:l^2\to l^2, x=(x_i)\mapsto(1-|x|, x_1, x_2, \cdots),$$

则易验证 $F\in C(B,B)$，$F$ 没有不动点.

以下结果给出应用 1.4.1 的一种常见形式.

**1.4.5 定理**(Leray-Schauder) 设 $F:X\to X$ 在有界集上为 SC，$\rho=\sup\{|x|:\exists\,\lambda\in(0,1)$，使 $x=\lambda Fx\}<\infty$，则 $\overline{B}_\rho(0)\cap\text{Fix}F\neq\varnothing$.

**证** $\forall\, n\geqslant1$，对 $F|\overline{B}(0,\rho+n^{-1})$ 应用 1.4.1 得出 $A_n\triangleq\overline{B}(0,\rho+n^{-1})\cap\text{Fix}F\neq\varnothing$. 因 $\text{Fix}F=(I-F)^{-1}(0)$，由 1.2.4 知 $A_n$ 紧，因此 $\overline{B}_\rho(0)\cap\text{Fix}F=\bigcap A_n\neq\varnothing$. □

在 1.4.1~1.4.3 及 1.4.5 的许多具体应用中，验证 $F$ 全连续是不成问题的. 如果 $F$ 非全连续而试图验证 $F\in \mathrm{SC}$，除了一些简单的特殊情况外，可能遇到计算非紧测度的困难. Mönch ([123]，1980)的两个结果以很独特的方式避开了这类困难.

**1.4.6 定理** 设 $D\subset X$ 闭凸，$F\in C(D,D)$ 满足条件

($\mathrm{H}_1$) 凡使 $\overline{C}=\overline{\mathrm{co}}(\{x_0\}\cup FC)$ 的可数集 $C\subset D$ 相对紧，其中 $x_0\in D$ 固定，则 $\mathrm{Fix}F\neq\varnothing$.

**证** 令 $A_0=\{x_0\}$，$A_{n+1}=\overline{\mathrm{co}}(\{x_0\}\cup FA_n)(n\geqslant 0)$，则 $\{A_n\}$ 是紧凸集的升列. 令 $A=\overline{\bigcup A_n}$，则 $A$ 闭凸，且 $FA\subset A=\overline{\mathrm{co}}(\{x_0\}\cup FA)$. 取可数集 $C_n\subset A_n$，使 $\overline{C_n}=A_n$，令 $C=\bigcup C_n$，则 $\overline{C}=A$. 于是

$$\overline{\mathrm{co}}(\{x_0\}\cup FC)\subset\overline{\mathrm{co}}(\{x_0\}\cup FA)\subset\overline{\mathrm{co}}(\{x_0\}\cup FC),$$

可见 $\overline{C}=\overline{\mathrm{co}}(\{x_0\}\cup FC)$. 由条件($\mathrm{H}_1$)，$\overline{C}$ 紧. 对 $F:A\to A$ 应用 1.4.3 得出 $\mathrm{Fix}F\neq\varnothing$. □

**1.4.7 定理** 设 $\Omega\subset X$ 是有界开集，$F\in C(\overline{\Omega},X)$，$x_0\in\Omega$，$(\lambda,x)\in J\times\partial\Omega\Rightarrow x-x_0\neq\lambda(Fx-x_0)$. 若 $F$ 满足条件

($\mathrm{H}_2$) 凡使 $C\subset\overline{\mathrm{co}}(\{x_0\}\cup FC)$ 的可数集 $C\subset\overline{\Omega}$ 相对紧，则 $\mathrm{Fix}F\neq\varnothing$.

我们略去这个定理的证明. 所宜注意者，1.4.6 与 1.4.7 分别蕴涵了 Darbo 定理与 1.4.1，而其本身的陈述完全不涉及非紧测度!

## §5 积分方程

现在转入应用前述结果的某些实例. 首先考虑如下积分方程：

$$x(t)=\int_T f(t,s,x(s))\mathrm{d}s\triangleq Fx(t),\qquad (1)$$

其中 $\mathrm{d}s$ 记 $\mathrm{d}\mu(s)$，$\mu$ 是紧 Hausdorff 空间 $T$ 上的正则测度，$\mu T>0$；$f:T\times T\times X\to X$. 有很多近代工作涉及用度与不动点方法研究方程(1)的可解性. 较精细的结果自然有赖于对 $T$ 与 $f$ 的较具体的

设定，无法在此深入讨论. 下面的结果有点一般化，但可初步阐释不动点方法的用法.

令 $Y=C(T,X)$，则求方程(1)的连续解相当于求 $F$ 在 $Y$ 中的不动点，$F$ 由(1)界定. 自然试图求助于上节的不动点定理. 对 $f$ 设定以下条件:

($H_1$) Caratheodory 条件：$\forall\ (t,x)\in T x X$，$f(t,\cdot,x)$ 为 $\mu$-可测；$\forall\ t\in T$，对 $\mu$-几乎所有 $s\in T$，$f(t,s,\cdot)$连续.

($H_2$) 积分连续性：$\forall\ t\in T$，有界集 $A\subset Y$，关于 $x\in A$ 一致地有

$$\lim_{\tau\to t}\int_T f(\tau,s,x(s))\mathrm{d}s=\int_T f(t,s,x(s))\mathrm{d}s.$$

($H_3$)增长性条件：$\exists\ a,b\in L^1(T,\mathrm{R}_+)$，使 $|f(t,s,x)|\leqslant a(s)|x|+b(s)(t,s\in T,x\in X)$，$|a|_1<1$.

($H_4$) 集压缩性条件：$\exists\ k\in[0,(2\mu T)^{-1})$，对任给有界集 $V\subset X$ 与 $t\in T$ 有 $\alpha(f(t,T,V))\leqslant k\alpha(V)$.

今从条件($H_1$)～($H_4$)推出以下结论：

1° $\forall\ x\in Y,t\in T$，条件($H_1$)保证 $f(t,\cdot,x(\cdot))$ $\mu$-可测. 由条件($H_3$)有

$$|f(t,s,x(s))|\leqslant a(s)|x|_0+b(s)\in L^1(T),\tag{2}$$

这推出 $Fx(t)$有定义. 条件($H_2$)表明 $t\mapsto Fx(t)$连续，即 $Fx\in Y$，这就得到算子 $F:Y\to Y$. 实际上，($H_2$)表明当 $A\subset Y$ 有界时 $FA$ 等度连续.

2° 设在 $Y$ 中 $x_n\to x$，则必在 $Y$ 中 $Fx_n\to Fx$. 否则有 $\varepsilon>0$，$t_n\in T$：$|Fx_n(t_n)-Fx(t_n)|\geqslant\varepsilon(n\geqslant1)$. 不妨设 $t_n\to t\in T$. 由$\{Fx_n\}$等度连续有$Fx_n(t_n)-Fx_n(t)\to0$. 由($H_1$)有$f(t,s,x_n(s))\to f(t,s,x(s))$，$\mu$-a.e.，于是用控制收敛定理(注意到(2))得出$Fx_n(t)\to Fx(t)$. 这就有

$$\begin{aligned}\varepsilon\leqslant&|Fx_n(t_n)-Fx_n(t)|+|Fx_n(t)-Fx(t)|\\&+|Fx(t)-Fx(t_n)|\to0\ (n\to\infty),\end{aligned}$$

得出矛盾．因此 $F\in C(Y,Y)$．

3° 由(2)推出 $|Fx|_0\leqslant|a|_1|x|_0+|b|_1(x\in Y)$．取 $\rho>0$ 充分大，令 $D=\{x\in Y:|x|_0\leqslant\rho\}$．$\forall\ x\in D$，由 $|a|_1<1$ 有 $|Fx|_0\leqslant\rho|a|_1+|b|_1\leqslant\rho$，因此 $FD\subset D$．

4° 任给 $A\subset D$，令 $V=A(T)$，显然 $V\subset X$ 有界．因 $FA$ 等度连续，故可用 1.1.4，结合 1.1.6 与 $(H_4)$ 有

$$\begin{aligned}\alpha(FA)&=\sup_{t\in T}\alpha((FA)(t))\\&=\sup_{t\in T}\alpha\left(\left\{\int_T f(t,s,x(s))\mathrm{d}s;x\in A\right\}\right)\\&\leqslant\sup_{t\in T}\mu T\cdot\alpha(\{f(t,s,x(s));s\in T,x\in A\})\\&\leqslant\sup_{t\in T}\mu T\cdot\alpha(f(t,T,V))\\&\leqslant\mu T\cdot k\alpha(V)\leqslant 2k\mu T\alpha(A),\end{aligned}$$

由此得 $\alpha(F)\leqslant 2k\mu T<1$．

综上所述，应用 Darbo 定理得到

**1.5.1 定理**　若条件 $(H_1)\sim(H_4)$ 满足，则方程(1)至少有一连续解．

如果以 1.4.6 取代 Darbo 定理，则条件 $(H_4)$ 可减弱为如下的

$(H'_4)$ $\exists\ k\in[0,(2\mu T)^{-1})$，$\forall\ t\in T$ 及 $\mu$-几乎所有 $s\in T$：$\alpha(f(t,s,\cdot))\leqslant k$．

事实上，若可数集 $C\subset D$（依前面的记号）满足

$$\bar{C}=\overline{\mathrm{co}}(\{0\}\cup FC),\tag{3}$$

则如下面所证的有 $\alpha(C)=0$．令 $Q=FC$，$A_t=\{f(t,\cdot,x(\cdot)):x\in C\}$．对任给 $t\in T$，由 $(H'_4)$ 有

$$\alpha(A_t(s))\leqslant\alpha(f(t,s,C(s)))\leqslant k\alpha(C(s)),\mu\text{-a. e.}.\tag{4}$$

由(3)易推出 $C(s)\subset\overline{\mathrm{co}}(\{0\}\cup Q(s))$，因此 $\alpha(C(s))\leqslant\alpha(Q(s))$，以此代入(4)得

$$\alpha(A_t(s))\leqslant k\alpha(Q(s)),\mu\text{-a. e.}.\tag{5}$$

由 $Q$ 等度连续推出 $t\mapsto\alpha(Q(t))$ 连续(1.1.4)；由(2)看出 $A_t$ 一致可积，于是用 1.1.8 及(5)有

$$\alpha(Q(t)) = \alpha\Big(\Big\{\int_T f(t,s,x(s))\mathrm{d}s ; x \in C\Big\}\Big)$$
$$= \alpha\Big(\Big\{\int_T y(s)\mathrm{d}s ; y \in A_t\Big\}\Big)$$
$$\leqslant 2\int_T \beta(A_t(s))\mathrm{d}s \leqslant 2k\int_T \alpha(Q(s))\mathrm{d}s.$$

上式两端对 $t$ 积分得出

$$\int_T \alpha(Q(t))\mathrm{d}t \leqslant 2k\mu T\int_T \alpha(Q(s))\mathrm{d}s.$$

因 $2k\mu T<1$，故必 $\int_T \alpha(Q(t))\mathrm{d}t=0$. 若对 $\mu$ 附加条件：任给非空开集 $W\subset T$ 有 $\mu W>0$（在具体问题中此要求不难满足），则 $\alpha(Q(t))\equiv 0$，于是由 1.1.4 有 $\alpha(Q)=0$；进而由(3)得 $\alpha(C)=0$.

表面上看，条件($H'_4$)与($H_4$)似乎差别甚小. 其实不然. 例如，若 $\dim X<\infty$，则条件($H_1$)推出 $\forall\ t\in T$，对 $\mu$-几乎所有 $s\in T$，$f(t,s,\cdot)$全连续，因而条件($H'_4$)自动满足；但即使 $V\subset X$ 有界，$f(t,T,V)$未必有界，更不必说($H_4$)满足. 因此，应用 1.4.6 常能带来本质的改进.

类似的方法可用到 Volterra 型积分方程：

$$x(t) = w(t) + \int_0^t f(t,s,x(s))\mathrm{d}s, \tag{6}$$

其中 $f:\mathrm{R}_+\times\mathrm{R}_+\times X\to X$，$w\in C(\mathrm{R}_+,X)$是给定的. 约定 $J_r=[0,r]$($r\geqslant 0$). 如同证 1.5.1 一样可得出：

**1.5.2 定理** 设 $f$ 满足条件：

($C_1$) $\forall\ t\in\mathrm{R}_+$：$f(t,\cdot,\cdot)|J_t\times X$ 满足 Caratheodory 条件；

($C_2$) $\forall\ t\in\mathrm{R}_+$，有界集 $V\subset X$，关于$(s,x)\in J_t\times V$ 一致地有 $\lim\limits_{\tau\to t} f(\tau,s,x)=f(t,s,x)$；

($C_3$) $\forall\ r>0$，$\exists\ a_r,b_r\in L^1(J_r,\mathrm{R}_+)$，$\forall\ (t,s,x)\in J_r\times J_t\times X$：$|f(t,s,x)|\leqslant a_r(s)|x|+b_r(s)$；

($C_4$) $\forall\ r>0$，$\exists\ k_r\geqslant 0$，$\forall\ t\in J_r$：$\alpha(f(t,s,\cdot))\leqslant k_r$，a.e.（$s\in J_t$），则$\exists\ r>0$，方程(6)在 $J_r$ 上有连续解.

**1.5.3 推论** 设 $f:\mathbf{R}_+\times X\to X$ 满足 Caratheodory 条件；$\forall r>0$，$\exists\ a_r,b_r\in L^1(J_r,\mathbf{R}_+)$，$\forall\ (t,x)\in J_r\times X$：$|f(t,x)|\leqslant a_r(t)|x|+b(t)$；对 $t\in J_r$ 有 $\alpha(f(t,\cdot))\leqslant k_r<\infty$，a.e.，则 IVP

$$x'=f(t,x),x(0)=\xi\in X$$

在某个 $J_r(r>0)$ 上有解.

1.5.3 由应用 1.5.2 于积分方程

$$x(t)=\xi+\int_0^t f(s,x(s))\mathrm{d}s$$

得出. 类似于 1.5.3 的结果可在[42]中找到.

## §6 ODE 的边值问题

ODE 的边值问题(BVP)无疑是一个有点"古老"的问题，然而它仍然占有数量可观的近代文献. 问题在于，迄今并未形成一套系统的方法，足以对付出现于 BVP 中的各种情况. 在这种情况下求助于度方法是很自然的. 近年来，应用度理论研究 ODE 的 BVP 的工作是如此之多(如看[41，113，115，123，130，159，177，188，189])，以至形成了某种势头. ODE 的 BVP 的提法与形式都较简单，确实便于用来检验多种分析工具的效力，尽管这种检验可能仅是初步的. 这多少解释了人们何以乐于对 BVP 使用度方法. 应当说，度方法的使用毕竟带来了一批新结果. 对这方面的大量工作做任何形式的概括都是困难的，本节仅能提供一个一般性的导引.

考虑 $n$ 阶 BVP：

$$\begin{cases}x^{(n)}-\sum_{i=1}^{n}A_i(t)x^{(i-1)}=f(t,\tilde{x})\ (t\in J); & (1)\\ Bx=0, & (2)\end{cases}$$

其中 $A_i\in L^1(J,L(\mathbf{R}^m))$，$f:J\times\mathbf{R}^{mn}\to\mathbf{R}^m$，$\tilde{x}=(x,x',\cdots,x^{(n-1)})$，$B$：$X\triangleq C^{n-1}(J,\mathbf{R}^m)\to\mathbf{R}^{mn}$. 约定 $X$ 中采用范数 $|x|_X=\max\limits_{1\leqslant i\leqslant n}|x^{(i-1)}|_0$，且设 $B\in L(X,\mathbf{R}^{mn})$(这一要求通常能被满足)，于是 $X_0\triangleq N(B)$是

$X$ 的闭子空间. 令 $Y=L^1(J,\mathrm{R}^m)$, $Z=\{x\in X_0: x^{(n-1)}\in \mathrm{AC}\}$, AC 记绝对连续函数类. 分别以 $Lx$ 与 $Fx$ 记方程(1)的两端, 则 $L:Z\to Y$ 是一线性算子; 对满足一定条件的 $f$ 有 $F:X\to Y$. 若 $x\in Z$ 满足"半线性"算子方程

$$Lx = Fx, \tag{3}$$

则称 $x$ 为 BVP(1)(2)的(广义)解. 形式上可将(3)转化为"不动点方程" $x=Tx$, $T=L^{-1}F$, 然后应用适当的不动点定理. 下面指明, 在适当条件下上述思路是可行的. 设定条件:

($H_1$) BVP"$Lx=0, Bx=0$"仅有零解;

($H_2$) $f:J\times\mathrm{R}^{mn}\to\mathrm{R}^m$ 满足 Caratheodory 条件;

($H_3$) $\forall\, r>0, \exists\, g_r\in L^1(J)$: $|f(t,z)|\leqslant g_r(t)$ ($t\in J, z\in\mathrm{R}^{mn}$, $|z|\leqslant r$).

**1.6.1 引理** ($H_1$)$\Rightarrow K\triangleq(L|Z)^{-1}\in L(Y,X)$.

**证** 显然($H_1$)$\Rightarrow L:Z\to Y$ 为单射. 令

$$M=\begin{pmatrix} 0 & I & & 0 \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & I \\ A_1 & A_2 & \cdots & A_n \end{pmatrix},$$

则 $M\in L^1(J,L(\mathrm{R}^{mn}))$. 由标准的线性微分方程结论, IVP"$W'=M(t)W, W(0)=\mathrm{id}$"有绝对连续的唯一解 $W:J\to\mathrm{GL}(\mathrm{R}^{mn})$. 令 $W=(W_{ki})_{n\times n}$, 则必 $W_{1i}\in C^{n-1}(J,L(\mathrm{R}^m))$, $W_{1i}^{(k-1)}=W_{ki}$, $LW_{1i}=0$ ($1\leqslant k,i\leqslant n$). 用常数变异法求得 IVP"$Lx=y\in Y, \tilde{x}(0)=0$"的唯一解

$$x(t)=\int_0^t W_{1n}(t-s)y(s)\mathrm{d}s\triangleq u(t,y). \tag{4}$$

任给 $z=(z_1,z_2,\cdots,z_n)\in\mathrm{R}^{mn}$, $y\in Y$, 定义

$$Az=\sum B(W_{1i}z_i); \tag{5}$$

$$x(t,z,y)=\sum W_{1i}(t)z_i+u(t,y), \tag{6}$$

则 $A\in L(\mathrm{R}^{mn})$, $Lx=y$, $Bx=Az+Bu(\cdot,y)$, 其中 $x=x(\cdot,z,$

$y$). 若令 $Az=0, y=0$, 则 $Lx=0, Bx=0$, 于是由($H_1$)得 $x=0$; 代入(6)后得

$$x(t,z,0)=\sum W_{1i}(t)z_i\equiv 0,$$

故 $$z_k=\sum_i W_{ki}(0)z_i=\sum_i W_{1i}^{(k-1)}(0)z_i=0\ (1\leqslant k\leqslant n),$$

因此 $z=0$. 可见 $A$ 可逆. 令 $\varphi(y)=(\varphi_i(y))=-A^{-1}Bu(\cdot,y)$, $Ky=x(\cdot,\varphi(y),y)$, 则 $LKy=y, BKy=0$, 结合(4)(6)有

$$Ky(t)=\sum W_{1i}(t)\varphi_i(y)+\int_0^t W_{1n}(t-s)y(s)\mathrm{d}s. \tag{7}$$

因此 $KY\subset Z, K=(L|Z)^{-1}\in L(Y,X)$. □

**1.6.2 引理** ($H_2$)($H_3$)$\Rightarrow F:X\to Y$ 为有界连续映射.

证明是直接的.

**1.6.3 引理** 若条件($H_1$)~($H_3$)满足, 则 $T\triangleq KF:X_0\to X_0$ 全连续, $K$ 依 1.6.1.

**证** 由 1.6.1 与 1.6.2, $T:X_0\to X_0$ 是有界连续映射. 任给有界集 $S\subset X_0$, 令 $Q=TS$, $r=\sup\limits_{x\in S}|x|_X$, $q=\sup\limits_{x\in Q}|x|_X$. 由条件($H_3$)有 $g\in L^1(J)$, $\forall\ x\in S, t\in J$: $|Fx(t)|\leqslant g(t)$. 若 $x\in S$, $z=Tx$, 则 $Lz=Fx$, 当 $0\leqslant t<\tau\leqslant 1$ 时

$$|z^{(i-1)}(\tau)-z^{(i-1)}(t)|=\left|\int_t^\tau z^{(i)}(s)\mathrm{d}s\right|\leqslant q(\tau-t)\ (1\leqslant i<n);$$

$$\begin{aligned}|z^{(n-1)}(\tau)-z^{(n-1)}(t)|&=\left|\int_t^\tau z^{(n)}(s)\mathrm{d}s\right|\\&=\left|\int_t^\tau\left[\sum A_i(s)z^{(i-1)}(s)+Fx(s)\right]\mathrm{d}s\right|\\&\leqslant q\sum\int_t^\tau|A_i(s)|\mathrm{d}s+\int_t^\tau g(s)\mathrm{d}s,\end{aligned}$$

可见 $\{z^{(i-1)}: z\in Q, 1\leqslant i\leqslant n\}$ 等度连续, 从而由 Arzela-Ascoli 定理得出 $Q$ 在 $X$ 中相对紧. 故 $T$ 全连续. □

结合 1.4.1 与 1.6.3 立即得到:

**1.6.4 定理** 设条件($H_1$)~($H_3$)满足. 若有有界开集 $\Omega\subset X_0$, 使得 $0\in\Omega$, $\forall\ \lambda\in(0,1)$, BVP

$$Lx = \lambda f(t, \tilde{x})(t \in J), Bx = 0 \tag{8$_\lambda$}$$

的任何解 $x \overline{\in} \partial\Omega$，则 BVP(1)(2)有解 $x \in \overline{\Omega}$.

应用 1.6.4 的关键在于构成满足所述要求的开集 $\Omega$，这恰是整个方法的最需技巧的部分，而且必定需要某些附加条件. 对 $\Omega$ 的最简单的选择是令 $\Omega = \{x \in X_0 : |x|_X < \rho\}$，其中 $\rho$ 充分大. 若 $(8)_\lambda$ 的解 $x$ 恒满足 $|x|_X < \rho$，则当然 $x \overline{\in} \partial\Omega$. 因此有

**1.6.5 推论** 若条件$(H_1)\sim(H_3)$满足，且

$$\sup\{|x|_X : x \in X_0, \exists\, \lambda \in (0,1), x \text{ 满足}(8)_\lambda\} < \infty, \tag{9}$$

则 BVP(1)(2)有解.

这样，判定 BVP(1)(2)是否可解，归结为对 BVP$(8)_\lambda$ 的解进行"先验估计"以验证(9). 这种估计强烈地依赖于 $L$，$B$ 与 $f$ 的具体性质，尤其依赖于对 $f$ 的一定增长性限制.

鉴于条件(9)并非总是容易验证，可考虑对 $\Omega$ 作更灵活的选择. 例如可取 $\Omega = \{x \in X_0 : |x|_X < \rho \text{ 且 } |x^{(n-2)}|_0 < r\}$，$0 < r < \rho$. 注意 $x \in \partial\Omega \Longleftrightarrow |x|_X \leqslant \rho$，$|x^{(n-2)}|_0 \leqslant r$ 且二者至少一个取等号. 类似于 1.6.5 有：

**1.6.6 推论** 设条件$(H_1)\sim(H_3)$满足，$\exists\, r > 0$，使得

$$\sup\{|x|_X : x \in X_0, \exists\, \lambda \in (0,1), \text{使 } x \text{ 满足 } (8)_\lambda \text{ 且 } |x^{(n-2)}|_0 \leqslant r\} < \infty; \tag{10}$$

当 $x$ 满足 $(8)_\lambda(\lambda \in (0,1))$ 时 $\max_J |x^{(n-2)}(t)| \neq r$，则 BVP(1)(2)有解.

试看两个具体例子.

**1.6.7 例** 考虑 2 阶 Dirichlet 问题

$$x'' = f(t, x, x')(t \in J), x(0) = x(1) = 0, \tag{11}$$

其中 $f: J \times \mathrm{R}^{2m} \to \mathrm{R}^m$ 满足 Caratheodory 条件及条件

$(H_4)$存在 $p, q, r \in L^1(J, \mathrm{R}_+)$，使得

$$|f(t, x, y)| \leqslant p(t)|x| + q(t)|y| + r(t)(t \in J, x, y \in \mathrm{R}^m); \tag{12}$$

$$|p|_1 + 2|q|_1 < 2. \tag{13}$$

令$X=C^1(J,\mathbf{R}^m)$，$|x|_X=\max\{|x|_0,|x'|_0\}$，$X_0=\{x\in X: x(0)=x(1)=0\}$. 今验证 1.6.5 之条件. 直接计算知 BVP“$x''=0, x(0)=x(1)=0$”仅有零解，即条件$(H_1)$满足. 实际上，对于本例可求得(7)的具体形式：

$$Ky(t)=\int_0^1 G(t,s)y(s)\mathrm{d}s, \tag{14}$$

其中

$$G(t,s)=\begin{cases} s(t-1), s\leqslant t; \\ t(s-1), s>t. \end{cases} \tag{15}$$

其次显然条件$(H_4)$蕴涵$(H_3)$. 设 $\lambda\in(0,1)$，$x\in X_0$ 满足

$$x''=\lambda f(t,x,x'). \tag{16}$$

令 $\mu=|x'|_2$. 由分部积分有

$$\mu^2=\int_0^1(x'(t),x'(t))\mathrm{d}t=(x(t),x'(t))\Big|_0^1-\int_0^1(x(t),x''(t))\mathrm{d}t$$

$$=-\int_0^1(x(t),x''(t))\mathrm{d}t\leqslant|x|_0|x''|_1. \tag{17}$$

$\forall\, t\in J$，由 $x(0)=x(1)=0$ 并用 Schwarz 不等式得

$$|x(t)|\leqslant\min\left\{\int_0^t|x'(s)|\mathrm{d}s,\int_t^1|x'(s)|\mathrm{d}s\right\}$$

$$\leqslant\min\{\sqrt{t},\sqrt{1-t}\}|x'|_2\leqslant\frac{\mu}{\sqrt{2}},$$

故得$|x|_0\leqslant\mu/\sqrt{2}$. 由 $x'=(Kx'')'$ 与(14)(15)有

$$|x'(t)|=\left|\int_0^1\partial_t G(t,s)x''(s)\mathrm{d}s\right|\leqslant|x''|_1,$$

这得$|x'|_0\leqslant|x''|_1$. 由条件$(H_4)$与(16)有

$$|x''|_1\leqslant|p|_1|x|_0+|q|_1|x'|_0+|r|_1$$

$$\leqslant\frac{\mu}{\sqrt{2}}|p|_1+|q|_1|x''|_1+|r|_1.$$

因(13)推出$|q|_1<1$，故可从上式解出

$$|x''|_1\leqslant(1-|q|_1)^{-1}\left(\frac{\mu}{\sqrt{2}}|p|_1+|r|_1\right). \tag{18}$$

以此代入(17)，可解出

$$\mu \leqslant \sqrt{2}|r|_1(2-|p|_1-2|q|_1)^{-1}.$$

这与$|x|_0 \leqslant \mu/\sqrt{2}$，$|x'|_0 \leqslant |x''|_1$ 及(18)一起表明，相应于(9)的条件满足，于是由 1.6.5 知 BVP(11)有解.

若改令 $Y=C(J,\mathbf{R}^n)$，$Z=X_0 \cap C^n$，则当 $f$ 连续且$(H_1)$满足时 1.6.1～1.6.6 仍然成立，其证明所需的修改是明显的.

**1.6.8 例** 考虑 2 阶周期 BVP：

$$\begin{cases} x''=f(t,x,x') \ (t \in J); \\ x(0)=x(1), x'(0)=x'(1), \end{cases} \tag{19}$$

其中 $f \in C(J \times \mathbf{R}^{2m}, \mathbf{R}^m)$，$f(0,\cdot,\cdot)=f(1,\cdot,\cdot)$. 假定 $f$ 满足

$(H_5)$ $\exists\, a,b,c,r>0$，使$|f(t,x,y)| \leqslant a|x|+b|y|+c$ $(t \in J$，$x,y \in \mathbf{R}^m)$，且当$|x|=r$，$(x,y)=0$ 时$(x,f(t,x,y)) \geqslant 0$.

设 $X$ 与$|\cdot|_X$仍如前例，$X_0=\{x \in X: x(0)=x(1), x'(0)=x'(1)\}$，$Z=X_0 \cap C^2$，$Y=C(J,\mathbf{R}^m)$，$Lx=x''-ax$，$Fx=f(t,x,x')-ax$. 不难验知 $K=(L|Z)^{-1} \in L(Y,X)$，记 $T=KF: X_0 \to X_0$. 设 $\lambda \in (0,1)$，$x=\lambda Tx$，$|x|_0 \leqslant r$. $x=\lambda Tx$ 意味着 $x \in X_0$ 且

$$x''=a(1-\lambda)x+\lambda f(t,x,x'). \tag{20}$$

令 $\mu=|x'|_2$，则如同 1.6.7 有 $\mu^2 \leqslant |x|_0|x''|_1$，因而由(20)、$(H_5)$有

$$\begin{aligned} \mu^2 &\leqslant r\int_0^1 [a(1-\lambda)|x(t)|+\lambda a|x(t)|+\lambda b|x'(t)|+\lambda c]\mathrm{d}t \\ &\leqslant ar^2+br\mu+cr, \end{aligned}$$

这推出 $\mu \leqslant \text{const}$. 这结合(20)与$(H_5)$得$|x''|_1 \leqslant \text{const}$. 设 $x=(x_i)$. 由 $x(0)=x(1)$有 $t_i \in (0,1): x'_i(t_i)=0$，于是

$$|x'_i(t)|=\left|\int_{t_i}^t x''_i(s)\mathrm{d}s\right| \leqslant |x''_i|_1 \ (1 \leqslant i \leqslant n),$$

这推出$|x'|_0 \leqslant \text{const}$，因此相应于(10)的条件满足.

其次设 $x$ 如上段，则$\max\limits_J |x(t)| \neq r$，否则$\exists\, t_0 \in J: |x(t_0)|=r$，$\varphi(t) \triangleq (x(t),x(t_0))$在$t=t_0$取极大. 因$\varphi \in C^2$，故$\varphi'(t_0)=(x'(t_0),x(t_0))=0$，$\varphi''(t_0) \leqslant 0$，于是由条件$(H_5)$有

$$0 \geqslant (x''(t_0), x(t_0))$$
$$= a(1-\lambda)r + \lambda(x(t_0), f(t_0, x(t_0), x'(t_0)))$$
$$\geqslant a(1-\lambda)r > 0,$$

得出矛盾. 因此由 1.6.6(注意前面的说明)得出 BVP(19)有解.

## §7 LS 度与 Brouwer 度

如在§3 中引入度时所预期的,在前几节中,我们仅仅依据度的基本性质,经不多的演绎步骤,即达到了度理论所能引出的颇为深入的结论,而且接触到某些有价值的应用. 迄今为止,完全未谈及度的具体构成,更不必说度的计算公式. 从逻辑上说,以基本性质$(D_1)$~$(D_3)$作为公理,由之出发应能展开整个度理论. 然而,这一展开过程只能"迂回地"完成,即一般理论的建立有赖于对 LS 度与 Brouwer 度这两种特殊情况的充分展开,否则不足以克服某些技术性障碍. 而且,对特殊情形的深入考察有助于达到对度的更好的理解. 不过,对 LS 度与 Brouwer 度的完全的讨论并非短短几页所能做到,我们仅希望本节起某种阐释性作用,因而并未顾及逻辑上完全的严格性.

**1.7.1 命题** 设 $\Omega\subset X$ 为有界开集,$F\in \mathrm{SC}(\overline{\Omega})$,$\partial\Omega\cap \mathrm{Fix}F=\varnothing$,则存在紧凸集 $K\subset X$ 与 $F|(\overline{\Omega}\cap K)$的扩张 $\widetilde{F}\in\mathscr{K}(\overline{\Omega})$,使得 $\mathrm{Fix}\widetilde{F}=\mathrm{Fix}F\subset K$,且

$$d(I-F,\Omega,0)=d(I-\widetilde{F},\Omega,0). \tag{1}$$

**证** 可设 $\mathrm{Fix}F\neq\varnothing$(否则取 $K=\varnothing$,$\widetilde{F}x\equiv x_0\overline{\in}\overline{\Omega}$ 好了). 令 $\beta=\{A\subset X: A$ 闭凸,$\mathrm{Fix}F\cup F(\overline{\Omega}\cap A)\subset A\}$,则易验证 $K\triangleq\bigcap\limits_{A\in\beta}A\in\beta$,$C\triangleq\overline{\mathrm{co}}F(\overline{\Omega}\cap K)\subset K$,$C\in\beta$,因此 $K=C$. 由 $\alpha(K)=\alpha(F(\overline{\Omega}\cap K))\leqslant\alpha(F)\alpha(K)$推出 $\alpha(K)=0$,故 $K$ 紧凸. 由 1.2.5,$F|(\overline{\Omega}\cap K)$有紧扩张 $\widetilde{F}$:$\widetilde{F}\overline{\Omega}\subset\mathrm{co}F(\overline{\Omega}\cap K)\subset K$. 显然 $\mathrm{Fix}\widetilde{F}=\mathrm{Fix}F\subset\overline{\Omega}\cap K$. 令 $K_n=\{x:\mathrm{d}(x,\overline{\Omega}\cap K)\geqslant 1/n\}$,$n$ 充分大,则由$(D_4)(D_5)$推出

$$d(I-F,\Omega,0)=d(I-F,\Omega\backslash K_n,0)$$
$$=d(I-\widetilde{F},\Omega\backslash K_n,0)=d(I-\widetilde{F},\Omega,0).$$ □

**注** 1.7.1 的意义在于,可以用一个 LS 度 $d(I-\widetilde{F},\Omega,0)$ 来提供有关 Fix$F$ 的信息,而不必 $F\in\mathscr{K}$. 因此可以说,原则上 LS 度是足够的!

**1.7.2 命题** 设 $\Omega\subset X$ 是有界开集,$F\in\mathscr{K}(\overline{\Omega})$,$\partial\Omega\cap\text{Fix}F=\varnothing$,则存在有限维子空间 $Y\subset X$ 与 $G\in C(\overline{\Omega},Y)$,使

$$d(I-F,\Omega,0)=d(I-G,Y\cap\Omega,0). \tag{2}$$

**证** 因 $(I-F)(\partial\Omega)$ 闭(1.2.4),故 $2\varepsilon\triangleq\inf\limits_{x\in\partial\Omega}|x-Fx|>0$. 由 1.2.6,存在有限维子空间 $Y\subset X$ 与 $G\in C(\overline{\Omega},Y)$,使 $|F-G|_0<\varepsilon$. 由$(D_5)(D_6)$得出(2). □

1.7.2 表明,LS 度可归结为 Brouwer 度,因而可以 Brouwer 度作为构造度理论的出发点. 由于在有限维空间中有非常成熟的拓扑与分析工具可用,Brouwer 度是较易把握与操作的.

**1.7.3 命题** 设 $\Omega\subset\mathbb{R}^n$ 是有界开集,$f\in C(\overline{\Omega},\mathbb{R}^n)$,$0\overline{\in}f(\partial\Omega)$,则存在 $g\in C^1(\overline{\Omega},\mathbb{R}^n)$,使得 0 是 $g$ 的正则值(即 $\forall\ x\in g^{-1}(0):g'(x)\in\text{GL}(\mathbb{R}^n)$),且 $d(f,\Omega,0)=d(g,\Omega,0)$.

**证** 显然 $2\varepsilon\triangleq d(0,f(\partial\Omega))>0$. 取 $h\in C^\infty(\overline{\Omega},\mathbb{R}^n)$,使 $|f-h|_0<\varepsilon$[75;8.3.6]. 由 Sard 定理[75;5.5.3],$h$ 有正则值 $y$: $|y|<\varepsilon$. 令 $g=h-y$,则 0 是 $g$ 的正则值,$|f-g|_0<2\varepsilon$,于是由$(D_5)$有 $d(f,\Omega,0)=d(g,\Omega,0)$. □

1.7.3 表明,对于 Brouwer 度可限于考虑足够"好"的函数,这就为建立某种计算公式开辟了道路.

**1.7.4 定理** 设 $\Omega\subset\mathbb{R}^n$ 是有界开集,$f\in C^1(\overline{\Omega},\mathbb{R}^n)$,$0\overline{\in}f(\partial\Omega)$,0 是 $f$ 的正则值,则

$$d(f,\Omega,0)=\sum_{f(x)=0}\text{sgn}\det f'(x). \tag{3}$$

**证** 任给 $x_0\in f^{-1}(0)$,必 $x_0\in\Omega$,且 $f'(x_0)\in\text{GL}(\mathbb{R}^n)$. 由反函数定理,$f$ 在 $x_0$ 邻近为局部 $C^1$ 同胚,因此 $x_0$ 是 $f$ 的孤立零点. 可

见 $f^{-1}(0)$ 必为有限集，设为 $\{x_0, x_1, \cdots, x_m\}$. 取充分小的 $\varepsilon>0$，由 $(D_2)$ 有

$$d(f,\Omega,0)=\sum d(f,B_\varepsilon(x_i),0). \tag{4}$$

由 1.3.3 有 $d(f,B_\varepsilon(x_i),0)=d(f'(x_0),B_\varepsilon(0),0)$，于是(3)由(4)及下面的线性结果(1.7.5)得出. □

**1.7.5 定理** 设 $A\in GL(R^n)$，$\varepsilon>0$，则

$$d(A,B_\varepsilon(0),0)=\operatorname{sgndet}A=(-1)^\beta, \tag{5}$$

$\beta$ 是 $A$ 的负特征值的代数重数之和.

**证** 不妨设 $\varepsilon=1$，$A$ 是 Jordan 标准形. 由 $(D_5)$，若 $A(\cdot)\in C([a,b],GL(R^n))$，则 $d(A(t),B_1(0),0)$ 与 $t$ 无关；显然 sgndet $A(t)$ 亦与 $t$ 无关. 基于此，证明归于将 $A$ 沿 $GL(R^n)$ 内的单参数曲线作连续变形. 首先将每个 Jordan 块 $\begin{pmatrix}\cdot & 1 & & \\ & \cdot & \ddots & \\ & & \ddots & 1\\ & & & \cdot\end{pmatrix}$ 中的 1 代以参数 $t$ 并令其递降至零，于是 $A$ 变为 $A_1\oplus A_2\oplus\cdots\oplus A_k$，$A_i$ 为对角形或形如 $\begin{pmatrix}\alpha & \beta\\ -\beta & \alpha\end{pmatrix}$，$\alpha^2+\beta^2\neq0$. 以 $\begin{pmatrix}r\cos t & r\sin t\\ -r\sin t & r\cos t\end{pmatrix}$ 替代 $\begin{pmatrix}\alpha & \beta\\ -\beta & \alpha\end{pmatrix}$ 并令 $t$ 递降至零得到 $\begin{pmatrix}r & 0\\ 0 & r\end{pmatrix}$. 于是可设 $A=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$，$0\neq\lambda_i\in R$；进而可设 $\lambda_i=\pm1$. 因 $\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$ 可代以 $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$，故只需考虑 $A=\operatorname{diag}(1,\cdots,1,\pm1)$. 若 $A=I$，则(5)由 $(D_1)$ 得出. 若 $A=\operatorname{diag}(1,\cdots,1,-1)$，则不妨设 $n=1$，从而 $A=-I$. 令 $\Omega=(-1,1)$，$\Omega_1=(-1,0)$，$\Omega_2=(0,1)$，$f(x)=|x|-1/2$. 用一同伦论证易得 $d(f,\Omega,0)=0$，于是由 $(D_2)(D_3)$ 有

$$\begin{aligned}0&=d(f,\Omega_1,0)+d(f,\Omega_2,0)\\ &=d\left(-I-\frac{1}{2},\Omega,0\right)+d\left(I-\frac{1}{2},\Omega,0\right)\\ &=d(-I,\Omega,0)+d(I,\Omega,0),\end{aligned}$$

由此得出 $d(-I,\Omega,0)=-1$，如所要证. □

1.7.4 与 1.7.5 不仅给出了 Brouwer 度的明显算式，而且也对"方程 $f(x)=0$ 的解的代数个数"赋予了准确的含义，至少对性质充分"正则"的 $f$ 是如此. 这样，Brouwer 度就成为一种易于理解与把握的概念.

在无穷维空间中不再有对应于(3)的公式. 至于(5)，倒是有一个优美的无穷维推广：

**1.7.6 定理**[43;Th. 9.10] 设 $T\in L(X)\cap SC$，$1\overline{\in}\sigma(T)$，则

$$d(I-T,B_\varepsilon(0),0)=(-1)^\beta,$$

其中 $\varepsilon>0$，$\beta$ 是 $T$ 在 $(1,\infty)$ 内的谱值的代数重数之和.

**注** 若 1.3.4 中条件$(H_2)$或$(H_3)$满足，则直接用 1.7.6 得出 $d(f,\Omega,0)=1$.

## §8 Borsuk 定理

任给对称集 $D\subset X$，设 $f:D\to Y$，称 $\mathrm{odd}f(x)\triangleq[f(x)-f(-x)]/2$ 为 $f$ 的"奇部"；$f$ 是奇函数 $\Longleftrightarrow \mathrm{odd}f=f$，$f$ 是偶函数 $\Longleftrightarrow \mathrm{odd}f=0$；$\mathrm{odd}f$ 必为奇函数.

**1.8.1 Borsuk 定理** 设 $\Omega\subset X$ 为有界对称开集，$0\in\Omega$，$F=I-f\in SC(\overline{\Omega})$. 若 $f$ 满足条件

(H) $\forall\ (t,x)\in J\times\partial\Omega: f(x)\neq tf(-x)$，

则 $d(f,\Omega,0)=$奇数. 特别，若 $f|\partial\Omega$ 是奇函数且 $0\overline{\in}f(\partial\Omega)$，则 $d(f,\Omega,0)=$奇数.

**证** 下面分几步相继归化到特殊情况.

首先令 $H(t,x)=(1+t)^{-1}[Fx-tF(-x)]$，则由 1.2.7 推出 $H:J\times\overline{\Omega}\to X$ 为 SC-同伦；条件(H)推出 $\partial\Omega\cap\mathrm{Fix}H_t=\varnothing$；$H_0=F$，$H_1=\mathrm{odd}F$. 因为以 $H_1$ 代 $F$，故下面不妨设 $F$ 是奇函数，且 $\partial\Omega\cap\mathrm{Fix}F=\varnothing$.

其次设 $K,\tilde{F}$ 依 1.7.1，则不难验知 $K=-K$，$F_1\triangleq\mathrm{odd}\tilde{F}\in\mathscr{K})(\overline{\Omega})$，$F_1|(\overline{\Omega}\cap K)=F|(\overline{\Omega}\cap K)$，$F_1\overline{\Omega}\subset K$. 因可以 $F_1$ 代 $F$，故

不妨设 $F\in\mathscr{K}(\overline{\Omega})$.

若在1.2.6之证中取$\{y_i\}$为对称集，则当$F$为奇函数时由§2(4)定义的 $G$ 亦为奇函数. 基于这一事实与$(D_6)$，以下不妨设 $X=\mathrm{R}^n$，即 $\Omega\subset\mathrm{R}^n$，$f\in C(\overline{\Omega},\mathrm{R}^n)$为奇函数，$0\overline{\in}f(\partial\Omega)$.

取$g\in C^\infty(\overline{\Omega},\mathrm{R}^n)$，使$|f-g|_0$充分小；取充分小的$\varepsilon>0$，使$\varepsilon\overline{\in}\sigma(g'(0))$；令 $h=\mathrm{odd}g-\varepsilon I$，则 $h\in C^1(\overline{\Omega},\mathrm{R}^n)$是奇函数，$\det h'(0)\neq0$，$|h-f|_0$ 充分小. 因可用 $h$ 代 $f$，故不妨设 $f\in C^\infty$，$\det f'(0)\neq0$.

令 $U_i=\{x=(x_1,x_2,\cdots,x_n)\in\Omega;x_i\neq0\}$，$\Omega_k=\bigcup_1^k U_i(1\leqslant k\leqslant n)$，$\varphi(t)=t^3$. 取 $f(x)/\varphi(x_1)$在 $\Omega_1$ 上的正则值 $y^1$，使$|y^1|$充分小，则 0 是奇函数$g_1(x)\triangleq f(x)-\varphi(x_1)y^1$ 在 $\Omega_1$ 上的正则值. 设 $k<n$，已作出 $C^\infty$奇函数 $g_k$，使$|f-g_k|_0$ 充分小，0 是 $g_k|\Omega_k$ 的正则值，取 $g_k(x)/\varphi(x_{k+1})$在 $U_{k+1}$上的正则值 $y^{k+1}$，使$|y^{k+1}|$充分小，则 0 是 $C^\infty$奇函数 $g_{k+1}(x)\triangleq g_k(x)-\varphi(x_{k+1})y^{k+1}$在 $\Omega_{k+1}$上的正则值，$|f-g_{k+1}|_0$ 充分小. 如此直至得出 $C^\infty$奇函数 $g_n$，$|f-g_n|_0$ 充分小，0 是 $g_n|\Omega_n$ 的正则值. 令 $g=g_n$，则 $g'(0)=g'_{n-1}(0)=\cdots=f'(0)$，而 $\Omega_n=\Omega\backslash\{0\}$，故 0 亦是 $g|\Omega$ 的正则值. 因 $g^{-1}(0)$是对称集，而 $g'$是偶函数，故结合$(D_5)$与 1.7.4 得出

$$d(f,\Omega,0)=d(g,\Omega,0)$$
$$=\mathrm{sgn}\det g'(0)+\sum_{0\neq x\in g^{-1}(0)}\mathrm{sgn}\ \det g'(x)=\text{奇数}. \qquad \square$$

Borsuk 定理的结论异常深刻，仅仅是它所提供信息的一部分，即“奇函数⇒奇数度⇒不动点存在”，就足以导致深刻的结论. 下面是某些典型的应用.

**1.8.2 定理**(Borsuk-Ulam)　设 $\Omega\subset X$ 为有界对称开集，$0\in\Omega$；$F=I-f\in\mathrm{SC}(\overline{\Omega})$，$f(\partial\Omega)$含于 $X$ 的某真子空间，则有 $x\in\partial\Omega$，使 $f(-x)=f(x)$.

**证**　若定理结论不真，则 $g\triangleq\mathrm{odd}f$ 在 $\partial\Omega$ 上无零点，而 $I-g=\mathrm{odd}F\in\mathrm{SC}(\overline{\Omega})$，于是由1.8.1有$d(g,\Omega,0)\neq0$. 另一方面，因

$g(\partial\Omega)$如同$f(\partial\Omega)$一样含于$X$的某真子空间，从而$X\backslash g(\partial\Omega)$连通，故由($D_7$)有$d(g,\Omega,0)=0$，得出矛盾. □

若$f\in C(S^n,R^n)$，则$f$有扩张$\tilde{f}\in C(R^{n+1},R^n)$(1.2.7)，于是由1.8.2得出：

**1.8.3 推论** 若$f\in C(S^n,R^n)$，则$\exists\ x\in S^n: f(-x)=f(x)$.

当$n=2$时，1.8.3的结论的一些有趣的地理学解释在数学史上颇引人注目.

**1.8.4 推论**(三明治问题) 设$A_i\subset R^n$，$\text{mes}A_i<\infty(1\leqslant i\leqslant n)$，则有超平面$P$同时分$A_i$为等测度的两半.

**证** 任给$x=(x_0,x_1,\cdots,x_n)\in S^n$，令$H_x=\{(y_1,y_2,\cdots,y_n)\in R^n: \sum_1^n x_iy_i>x_0\}$，$f_i(x)=\text{mes}(A_i\cap H_x)$，则$f=(f_1,f_2,\cdots,f_n)\in C(S^n,R^n)$，于是$\exists\ x\in S^n: f(-x)=f(x)$，$P=\partial H_x$即为所求超平面. □

**1.8.5 推论**(Borsuk, 1933) 设$\Omega\subset R^n$是有界对称开集，$0\in\Omega$，$\partial\Omega=\bigcup_1^m A_i$，$A_i$是非空闭集，$A_i\cap(-A_i)=\varnothing(1\leqslant i\leqslant m)$，则$m>n$.

**证** 设$m\leqslant n$. 令$f_i(x)=d(x,A_i)$，则$f=(f_1,f_2,\cdots,f_{m-1})\in C(\bar{\Omega},R^{m-1})$，于是由1.8.2有$x\in\partial\Omega: f(x)=f(-x)$. 若$\exists\ i<m: x\in A_i$，则$-x\bar{\in} A_i$，这与$f_i(x)=f_i(-x)=0$矛盾. 若$x\in A_m$，则必$\exists\ i<m: -x\in A_i$，亦将导致矛盾. □

**1.8.6 推论** 设$\Omega$如1.8.5，$f\in C(\partial\Omega,R^n)$是奇函数，$f(\partial\Omega)$含于$R^n$的某真子空间，则$f^{-1}(0)\neq\varnothing$.

**证** 由1.8.2(参考1.8.3)，$\exists\ x\in\partial\Omega: f(x)=f(-x)$；因$f$是奇函数，故$f(x)=-f(x)$，因此$f(x)=0$. □

**1.8.7 开映射定理** 设$U\subset X$是开集，$F=I-f: U\to X$局部为集压缩映射，$f$为局部单射，则$f$是开映射.

**证** $\forall\ x_0\in U$，今证$f(x_0)\in(fU)^\circ$. 不妨设$f(x_0)=x_0=0$. 取$r>0$充分小，令$\Omega=B_r(0)$，使$F|\bar{\Omega}\in SC$且$f|\bar{\Omega}$为单射. 令$H(t,x)=F(tx)-F((t-1)x)$，则$H_1=F$，$H_{1/2}$是奇函数，$\partial\Omega\cap\text{Fix}H_t=$

$\varnothing(t\in J)$，且不难验证 $H:J\times\overline{\Omega}\to X$ 为 SC-同伦. 于是对邻近 0 的 $y$ 有

$$d(f,\Omega,y)=d(I-H_{1/2},\Omega,0)\neq 0,$$

从而 $y\in f(\Omega)$. 这表明 $0\in(fU)^{\circ}$，如所要证. □

在 1.8.7 的条件下，若 $f$ 是单射，则 $f:U\to fU$ 是同胚；若 $fU$ 闭，则必 $fU=X$. 利用这些事实容易建立：

**1.8.8 满射定理** 设 $F=I-f:X\to X$ 在有界集上为 SC，则以下每个条件蕴涵 $fX=X$：(i) $f$ 为局部单射且弱强制（即 $|x|\to\infty\Rightarrow|f(x)|\to\infty$）；(ii) $\overline{\lim\limits_{|x|\to\infty}}|Fx|/|x|<1$；(iii) $\exists\ \lambda>0,\forall\ x,y\in X$：$|f(x)-f(y)|\geqslant\lambda|x-y|$；(iv) $X$ 是 Hilbert 空间，$f$ 强制，即 $|x|\to\infty\Rightarrow(f(x),x)/|x|\to\infty$.

**证** 若条件(i)满足，$f(x_n)\to y$，则 $A\triangleq\{x_n\}$ 有界，从而 $f\overline{A}$ 闭(1.2.4)，于是 $y\in f\overline{A}$. 这表明 $fX$ 闭，因此由 1.8.7 有 $fX=X$. 其次显然(iii)⇒(i).

若条件(ii)满足，则对任给 $y\in X$，当 $\rho>0$ 充分大时有 $|Fx|<|x-y|$ $(\forall\ x\in\partial B_\rho(0))$，于是由($D_5$)有

$$d(f,B_\rho(0),y)=d(I,B_\rho(0),y)=1,\tag{1}$$

因而 $y\in fX$. 这表明 $fX=X$.

若条件(iv)满足，$t\in J$，则

$$\begin{aligned}|tx+t'f(x)|&\geqslant(tx+t'f(x),x/|x|)\\&=t|x|+t'(f(x),x)/|x|\to\infty\ (|x|\to\infty).\end{aligned}$$

给定 $y\in X$，当 $\rho$ 充分大时 $|x|=\rho\Rightarrow y\overline{\in}[x,f(x)]$，于是同样由($D_5$)得出(1). □

Borsuk 定理亦可用于微分方程问题，下面举一简单例子. 考虑 BVP

$$\begin{cases}x^{(n)}-\sum\limits_{i=1}^{n}A_i(t)x^{(i-1)}=f(t,\tilde{x})\ (t\in J);&(2)\\ B\tilde{x}=0,&(3)\end{cases}$$

其中 $A_i\in L^1(J,L(\mathrm{R}^m))\ (1\leqslant i\leqslant n)$，$f:J\times\mathrm{R}^{mn}\to\mathrm{R}^m$ 满足 §6 中条

件$(H_2)(H_3)$，$\widetilde{x}=(x,x',\cdots,x^{(n-1)})$，$B\in L(X,R^{mn})$，$X=C(J,R^{mn})$. 以 $Lx$ 记(2)之左端，设定以下条件：

$(H_1)$　$\exists\ \beta>0,\varphi\in L^1(J),\forall\ \lambda\in[0,1]$，BVP

$$Lx=f_\lambda(t,\widetilde{x})(t\in J),B\widetilde{x}=0 \tag{4}_\lambda$$

的解 $x$ 满足 $|f(t,\widetilde{x}(t))|\leqslant\beta\varphi(t)|\widetilde{x}(t)|$ 时必为零解，其中

$$f_\lambda(t,z)=(1+\lambda)^{-1}[f(t,z)-\lambda f(t,-z)]. \tag{5}$$

$(H_2)$　$\exists\ r>0,\forall\ (t,z)\in J\times R^{mn}$：$|z|\geqslant r\Rightarrow|f(t,z)|\leqslant\beta\varphi(t)\cdot|z|$，其中 $\beta,\varphi$ 依$(H_1)$.

**1.8.9 定理**　若条件$(H_1)(H_2)$满足，则 BVP(2)(3)有解.

**证**　如所熟知，$x$ 是 BVP(2)(3)的解$\Longleftrightarrow z=\widetilde{x}$ 是

$$z'=M(t)z+g(t,z)(t\in J),Bz=0 \tag{6}$$

的解，(5)中 $M$ 依 1.6.1 之证，$g=(0,\cdots,0,f)$：$J\times R^{mn}\to R^{mn}$. 任给 $z\in X$，定义

$$Tz(t)=z(0)+Bz+\int_0^t[M(s)z(s)+g(s,z(s))]ds. \tag{7}$$

由一标准的论证知 $T:X\to X$ 全连续. 显然 $z\in X$ 是 BVP(6)的解$\Longleftrightarrow z=Tz$. 取充分大的 $\rho>0$，令 $\Omega=\{z\in X:|z|_0<\rho\}$，只需验证 $T:\overline{\Omega}\to X$ 满足 1.8.1 中条件(H).

$T$ 不满足(H)意味着存在$(\lambda,z)\in J\times\partial\Omega$，使得

$$z=(1+\lambda)^{-1}[Tz-\lambda T(-z)]. \tag{8}$$

可设 $\lambda>0$(否则已有 $z=Tz$). 将(7)代入(8)得

$$z(t)=z(0)+Bz+\int_0^t[M(s)z(s)+g_\lambda(s,z(s))]ds, \tag{9}$$

其中 $g_\lambda=(0,\cdots,0,f_\lambda)$. (9)表明 $z$ 满足 BVP：

$$z'=M(t)z+g_\lambda(t,z)(t\in J),Bz=0. \tag{10}$$

这又推出存在 $x\in C^{n-1}(J,R^m)$：$z=\widetilde{x}$，$x$ 满足 BVP$(4)_\lambda$.

若$\forall\ t\in J$：$|z(t)|\geqslant r$，$r$ 依条件$(H_2)$，则由$(H_2)$有

$$|f(t,z(t))|\leqslant\beta\varphi(t)|z(t)|(\forall\ t\in J),$$

于是由条件$(H_1)$推出 $x(t)\equiv0$，得出矛盾. 因此必有 $t_0\in J$：$|z(t_0)|$

$<r$. 若 $t_0<1$，则 $\forall\ t\in[t_0,1]$ 有

$$\begin{aligned}|z(t)| &= \left|z(t_0)+\int_{t_0}^{t}[M(s)z(s)+g_\lambda(s,z(s))]\mathrm{d}s\right| \\ &< r+\int_{t_0}^{t}[k|M(s)||z(s)|+|g_\lambda(s,z(s))|]\mathrm{d}s,\end{aligned} \tag{11}$$

$k>0$ 是某个常数. 取 $h\in L^1(J,\mathrm{R}_+)$，使当 $(t,z)\in J\times\mathrm{R}^{mn}$，$|z|\leqslant r$ 时 $|f(t,z)|\leqslant h(t)$. 这结合条件$(\mathrm{H}_2)$得出

$$|f(t,z(t))|\leqslant\beta\varphi(t)|z(t)|+h(t)\quad(\forall\ t\in J).$$

由此显然推出

$$|g_\lambda(t,z(t))|\leqslant\beta\varphi(t)|z(t)|+h(t)\quad(\forall\ t\in J).$$

以此代入(11)得

$$|z(t)|<r+|h|_1+\int_{t_0}^{t}\omega(s)|z(s)|\mathrm{d}s,$$

其中 $\omega(s)=k|M(s)|+\beta\varphi(s)$. 于是用 Gronwall 不等式得

$$|z(t)|\leqslant(r+|h|_1)\exp\left(\int_0^1\omega(s)\mathrm{d}s\right)\triangleq R\ (t_0\leqslant t\leqslant 1),$$

$R$ 与 $\rho$ 无关. 若 $t_0>0$，同理可证当 $0\leqslant t<t_0$ 时，$|z(t)|\leqslant R$，故 $|z|_0\leqslant R$. 因可设 $R<\rho$，故 $z\in\partial\Omega$，得出矛盾. □

类似于 1.8.9 的结果可在[188]中找到.

# 第二章　锥与正不动点

本章讨论有序Banach空间中的非线性映射及其不动点，其基本思想一方面可溯源于经典分析中关于单调函数的熟知结果；另一方面基于度理论的某种推广，后者导致关于“正不动点”的一系列结果，这些结果今天已广泛应用于各类非线性方程正解问题的研究.

## §1　锥

**2.1.1 定义**　设 $K\subset X$. 若 $\forall\ \lambda>0: \lambda K\subset K$，则称 $K$ 为锥；若 $K$ 是闭凸锥且 $K\cap(-K)=\{0\}\neq K$，则称 $K$ 为序锥.

锥是要求甚少而作用甚大的数学概念之一. 本书在两条途径上使用锥概念：其一是用作描述某些数学对象的“方向锥”，第四、五两章将予考虑；其二是用于定义序关系，这是本章的课题.

任何非空闭凸锥 $K\subset X$ 在 $X$ 中导入一拟序 $\leqslant$（必要指明 $K$ 时记作 $\leqslant_K$）：$x\leqslant y\Longleftrightarrow y-x\in K$. $\leqslant$ 具有通常的性质：若 $x\leqslant y, a\leqslant b, \lambda\in\mathrm{R}_+$，则 $x+a\leqslant y+b, \lambda x\leqslant\lambda y$；若 $x_n\leqslant y_n, x_n\to x, y_n\to y$，则 $x\leqslant y$. 若 $K$ 为序锥，则 $\leqslant$ 为半序，此时说 $X$（或写作 $(X,\leqslant)$）是有序Banach 空间. 为记号简便，本书约定：当写出 $X_+$ 时意味着 $X_+\subset X$ 为非空闭凸锥，在 $X$ 中使用由 $X_+$ 导入的序 $\leqslant$；$x<y\Longleftrightarrow x\leqslant y\neq x$，$x\ll y\Longleftrightarrow y-x\in X^\circ_+$；$[a,b]=\{x\in X: a\leqslant x\leqslant b\}$（序区间），$A\subset X$ 序有界 $\Longleftrightarrow\exists\ a,b\in X: A\subset[a,b]$；$A\leqslant B\Longleftrightarrow\forall\ a\in A, b\in B: a\leqslant b$；序列 $\{x_n\}\subset X$ 单调增及函数 $f: D\subset X\to Y$ 单调增的含义是自明的.

本节设 $X_+$ 是一序锥，它导入序 $\leqslant$.

**2.1.2 定义**　若 $X=\mathrm{span}X_+(=X_+-X_+)$，则说 $X_+$ 是再生

的；若 $X=\overline{\mathrm{span}X_+}$，则说 $X_+$是完全的；若 $X$ 中的序有界序列有界，则说 $X_+$是正规的；若 $X$ 中的序有界增序列收敛，则说 $X_+$是正则的；若 $X$ 中的有界增序列收敛，则说 $X_+$是全正则的.

**2.1.3 定理** $X^\circ_+\neq\varnothing\Rightarrow X_+$是再生的$\Longleftrightarrow\exists\ \delta>0:B_\delta(0)\subset C\triangleq(X_+\cap B_1(0))-(X_+\cap B_1(0))$.

**证** 若 $X^\circ_+\neq\varnothing$，则$\exists\ x_0\in X,r,\varepsilon>0:B_\varepsilon(x_0)\subset X_+\cap B_r(0)$. 令 $\delta=\varepsilon/r$，则 $B_\delta(0)\subset C$.

若 $B_\delta(0)\subset C$，则 $X=\bigcup_1^\infty nB_\delta(0)=\bigcup_1^\infty nC=X_+-X_+$.

若 $X=X_+-X_+$，则 $X=\bigcup_1^\infty n\overline{C}$，于是由 Baire 定理有 $\overline{C}^\circ\neq\varnothing$. 由 $\overline{C}^\circ$对称凸推出 $0\in\overline{C}^\circ$，于是$\exists\ \delta>0:B_{2\delta}(0)\subset\overline{C}$，今证 $B_\delta(0)\subset C$：$\forall\ x\in B_\delta(0)$，取 $x_1\in C\cap B_\delta(2x)$，$x_2\in C\cap B_\delta(4x-2x_1)$，$x_3\in C\cap B_\delta(8x-4x_1-2x_2)$，…，如此得 $\{x_n\}\subset C$，$x=\sum_1^\infty 2^{-n}x_n\in C$. □

"$X^\circ_+\neq\varnothing$"是一个不常有但极有用的性质，它反映了 $X_+$的某种"丰满性"，且与"再生性"、"完全性"一起构成"丰满程度"递降的三个层次.

**2.1.4 定理** 以下条件互相等价：(i)$X_+$是正规的；(ii)$\inf\{|x+y|:x,y\in X_+\cap\partial B\}>0$，$B=B_1(0)$；(iii)$\exists\ \beta>0:0\leqslant x\leqslant y\Rightarrow|x|\leqslant\beta|y|$；(iv)$(B+X_+)\cap(B-X_+)$有界.

**证** (i)⇒(ii). 若(ii)不真，则$\exists\ x_n,y_n\in X_+\cap\partial B:|x_n+y_n|<n^{-3}(n\geqslant1)$，于是 $0\leqslant nx_n\leqslant\sum_1^\infty k(x_k+y_k)$，$X_+$ 必非正规.

(ii)⇒(iii). 若(iii)不真，则$\exists\ x_n,y_n:0\leqslant x_n\leqslant y_n$，$|x_n|>n|y_n|>0$. 令 $x'_n=x_n/|x_n|$，$y'_n$，$z'_n$ 仿此，$z_n=y'_n-nx'_n$，则 $x'_n,z'_n\in X_+\cap\partial B$，$n-1\leqslant|z_n|\leqslant n+1$，$|x'_n+z'_n|\leqslant2/(n-1)$，(ii)不能成立.

(iii)⇒(iv). $\forall\ a,b\in B$，$x,y\in X_+$，当 $a+x=b-y$ 时有 $0\leqslant x\leqslant b-a$，于是从(iii)推出 $|x|\leqslant\beta|b-a|\leqslant2\beta$，从而 $|a+x|\leqslant2\beta+1$. 因此$(B+X_+)\cap(B-X_+)\subset\overline{B}_{2\beta+1}(0)$.

(iv)⇒(i). 设 $a\leqslant x_n\leqslant b$，可设 $a,b\in B$，于是 $x_n=a+(x_n-a)=b-(b-x_n)\in(B+X_+)\cap(B-X_+)$，故(iv)⇒$\{x_n\}$有界. □

正规性通常在以下形式下使用：设已给 $x$ 的"序估计"$a\leqslant x\leqslant b$，则 $\theta\leqslant x-a\leqslant b-a$，于是依 2.1.4(iii) 得出一"模估计"：$|x-a|\leqslant\beta|b-a|$.

**2.1.5 命题** 全正则⇒正则⇒正规.

**证** 若 $X_+$ 非正规，则有 $x_n,y_n\in X_+\cap\partial B_1(0)$：$|x_n+y_n|<n^{-2}$. 令 $z_{2n}=\sum_1^n(x_k+y_k)$，$z_{2n+1}=z_{2n}+x_{n+1}$，则 $\{z_n\}$ 是有界且序有界的发散增序列，因此 $X_+$ 非正则亦非全正则. 若 $X_+$ 全正则，$\{x_n\}\subset X$ 是增序列且序有界，则因 $X_+$ 正规 $\{x_n\}$ 必有界，从而收敛，于是 $X_+$ 正则. □

"有界增序列收敛"是熟知的实分析结论，这一结论在一般有序 Banach 空间中只是有条件地成立，而正则与全正则性正表达了这类条件.

**2.1.6 例** 1° 本书中如无特别声明，在 $\mathrm{R}^n$ 中总采用由锥 $\mathrm{R}_+^n=(\mathrm{R}_+)^n$ 导入的"标准向量序". $\mathrm{R}_+^n$ 内部非空且全正则. 实际上，不难证明 $\mathrm{R}^n$ 中任何序锥是全正则的，更是正规的. 因此，在有限维空间中不会有否定 2.1.4 之条件(i)～(iv)的直觉经验.

2° 设 $(\Omega,\mu)$ 是任何测度空间，$X=L^p(\Omega,\mathrm{R})$，$1\leqslant p<\infty$，则 $X_+\triangleq L^p(\Omega,\mathrm{R}_+)$ 是 $X$ 中的序锥，$X^\circ_+=\varnothing$，但 $X_+$ 是再生的：每个 $x\in X$ 有分解 $x=x^+-x^-$. 其次，不难用 Levi 定理推出 $X_+$ 是全正则的.

3° 设 $X=c_0$（收敛于零的实数列空间），则 $X_+\triangleq\{(x_i)\in c_0:x_i\geqslant0(\forall\ i\geqslant1)\}$ 是正则锥. $X_+$ 非全正则：令 $x_n=(\overbrace{1,\cdots,1}^{n\text{个}},0,0,\cdots)$，则 $\{x_n\}\subset c_0$ 是有界发散增序列.

4° 设 $X=C(T,\mathrm{R})$，$T$ 是紧 $T_2$ 空间，则 $X_+\triangleq C(T,\mathrm{R}_+)$ 是正规锥. $X_+$ 未必正则. 例如 $C(J,\mathrm{R}_+)$ 即非正则：设 $x_n(t)=1-t^n(t\in J)$，则 $\{x_n\}$ 是 $C(J,\mathrm{R})$ 中的序有界发散增序列.

5° 设 $X=C^1(J,\mathrm{R})$，在 $X$ 中采用范数 $|x|_X=|x|_0+|x'|_0$，则 $X_+=X\cap C(J,\mathrm{R}_+)$ 是 $X$ 中非正规的锥：$x_n=\sin nt(n\geqslant1)$ 在 $X$ 中

序有界但$|x_n|_X\to\infty(n\to\infty)$. 其次注意$X^\circ_+=X\cap C(J,(0,\infty))\neq\varnothing$. 此例与例2°一起说明:序锥内部非空与正规性之间没有必然联系.

从应用角度考虑,“$X^\circ_+=\varnothing$”与“$X_+$非正规”都是引起困难的“病态”. 然而,在无限维空间中,许多自然地定义的锥难免这种病态. 在这种情况下,有一个对空间与锥同时进行改造使之消除病态的标准方法. 设$K$是$X$中的任何序锥. 任取$0\neq e\in K$,令$X_e=\mathrm{R}_+[-e,e]$,$[-e,e]$是由$\leqslant_X$决定的序区间;在$X_e$中采用范数$|x|_e=\inf\{\lambda>0:x\in\lambda[-e,e]\}$,则易验知$(X_e,|\cdot|_e)$是一赋范空间,且$K_e\triangleq X_e\cap K=\mathrm{R}_+[0,e]$是$X_e$中的序锥. 如此构成的$K_e$有较好的性质.

**2.1.7 定理** $K_e$有以下性质:(i)$K_e$在$X_e$中的内部(下面记作$\mathrm{int}K_e$)非空,事实上,$\forall\ \lambda,\mu>0:[\lambda e,\mu e]\subset\mathrm{int}K_e$. (ii)$K_e$在$X_e$中正规. (iii)若$K$正规,则$(X_e,|\cdot|_e)$完备且它连续地嵌入$X$. (iv)若$e\in K^\circ$,则$X=X_e$,且当$K$正规时$|\cdot|_e$是$|\cdot|$的等价范数.

**证** (i)设$\lambda,\mu>0$,$\lambda e\leqslant x\leqslant\mu e$,$|x-y|_e<\varepsilon$,$\varepsilon$充分小,则$0\leqslant(\lambda-\varepsilon)e\leqslant y\leqslant(\mu+\varepsilon)e$,可见$y\in K_e$,因此$x\in\mathrm{int}K_e$.

(ii)以$\leqslant_e$记$K_e$在$X_e$中导入的序. 若$0\leqslant_e x\leqslant_e y$,则$0\leqslant x\leqslant y\leqslant|y|_e e$,这推出$|x|_e\leqslant|y|_e$,可见$K_e$正规.

(iii)若$K$正规,则$\exists\ r>0:[-e,e]\subset B_r(0)$,于是$\forall\ x\in X_e:|x|\leqslant r|x|_e$,可见$X_e\subset X$是连续嵌入. 若$|x_m-x_n|_e\to0(m,n\to\infty)$,则$|x_m-x_n|\to0(m,n\to\infty)$,于是在$X$中$x_n\to x$. $\forall\ \varepsilon>0$,当$m,n$充分大时有$x_n-x_m\in\varepsilon[-e,e]$;令$m\to\infty$得$x_n-x\in\varepsilon[-e,e]$,这表明在$X_e$中$x_n\to x$. 可见$X_e$完备.

(iv)设$B_\delta(e)\subset K$,$\delta>0$,则$B_\delta(0)\subset X_e$,因此$X=X_e$. 若$0\neq x\in X$,则$e\pm\delta x/|x|\geqslant0$,这推出$|x|_e\leqslant|x|/\delta$. 若$K$正规,则证(iii)时已得$|x|\leqslant r|x|_e$,因此$|\cdot|_e$与$|\cdot|$等价. □

## §2 对偶锥

首先约定本书将多次使用的某些缩记号. 任给 $A\subset X, U\subset X^*$, 令

$$\langle U, A\rangle = \{\langle u, a\rangle : u\in U, a\in A\}; \qquad (1)$$

$\langle U, a\rangle$ 与 $\langle u, A\rangle(a\in X, u\in X^*)$ 的意义自明. 记号 $u(A)$ 与 $\langle u, A\rangle$ 将同等地使用. 约定

$$A\leqslant B \Longleftrightarrow \forall\, a\in A, b\in B: a\leqslant b,$$

只要右端之 $a\leqslant b$ 有定义; $A<B, A\leqslant b$ 等仿此.

以下设 $X$ 中已由锥 $X_+$ 导入序 $\leqslant$.

**2.2.1 定义** 任给 $A\subset X$, 称 $A^*\triangleq\{u\in X^*: u(A)\geqslant 0\}$ 为 $A$ 的对偶锥; 称 $\mathrm{cone}A\triangleq\bigcup_{\lambda>0}\lambda A$ 为 $A$ 生成的锥, 约定 $\overline{\mathrm{cone}}A=\overline{\mathrm{cone}A}$. 称每个 $u\in X_+^*$ 为 $X$ 上的正线性泛函; 当 $\langle u, X_+\backslash\{0\}\rangle>0$ 时说 $u$ 是严格正的.

任给 $A\subset X$, 容易验证 $A^*$ 为 $X^*$ 中的弱*闭凸锥, $\overline{\mathrm{cone}\ \mathrm{co}A}=\overline{\mathrm{co}}\ \overline{\mathrm{cone}}A$, $A^*=(\overline{\mathrm{cone}}\ \overline{\mathrm{co}}A)^*$; 若 $A$ 是子空间, 则 $A^*=A^\perp$(此时须注意区别"对偶锥 $A^*$"与"对偶空间 $A^*$"!).

**2.2.2命题** 设$K\subset X$是非空锥. (i)$u\in K^*\Rightarrow\inf u(K)=0$, $u\in X^*\backslash K^*\Rightarrow\inf u(K)=-\infty$; 因此 $u(K)\geqslant r>-\infty\Rightarrow u\in K^*$ 且 $r\leqslant 0$. (ii)若$x\in K^\circ$, $0\neq u\in K^*$, 则$u(x)>0$; $u\in K^*$ 且 $|u|\to\infty\Rightarrow u(x)\to\infty$. (iii)若 $X^\circ_+\neq\varnothing$, $x\in\partial X_+$, 则$\exists\ u\in X_+^*\backslash\{0\}: u(x)=0$. (iv)若 $X$ 可分, 则存在严格正的 $u\in X_+^*$.

**证** (i)是明显的.

(ii)设$B_\delta(x)\subset K$, $\delta>0$, 则$0\leqslant\langle u, B_\delta(x)\rangle=u(x)-\langle u, B_\delta(0)\rangle$, 于是 $u(x)\geqslant\sup\langle u, B_\delta(0)\rangle=\delta|u|$, 由此得出所要证.

(iii)由分离定理有 $0\neq u\in X^*: u(x)<u(X^\circ_+)$, 于是 $u\in(X^\circ_+)^*=X_+^*$(注意 $X_+\subset\overline{X^\circ_+}$), $u(x)\leqslant 0\leqslant u(x)$, 从而 $u(x)=0$.

(iv) $X$ 可分推出 $X^*$ 弱*可分，于是有序列 $\{u_n\}$ 在 $X_+^* \cap B_1(0)$ 中弱*稠密，可验知 $u \triangleq \sum n^{-2}u_n$ 为严格正. □

以下命题表明，$X_+$ 与 $X_+^{**}(=(X_+^*)^*)$ 的关系颇类似于 $X$ 与 $X^{**}$ 的关系.

**2.2.3 命题** $X_+=X\cap X_+^{**}$；$X$ 自反 $\Rightarrow X_+=X_+^{**}$.

**证** 显然 $X_+\subset X_+^{**}$. $\forall\, x\in X\backslash X_+$，由分离定理有 $u\in X^*$：$u(x)<u(X_+)$，于是由 2.2.2(i) 有 $u\in X_+^*$，$u(x)<0$，可见 $x\overline{\in} X_+^{**}$. 故得 $X\cap X_+^{**}\subset X_+$. □

**注** 等式 $X_+=X\cap X_+^{**}$ 意味着，$\forall\, x\in X$：$x\geqslant 0\Longleftrightarrow\langle X_+^*,x\rangle\geqslant 0$. 其次，结合 2.2.3 与 2.2.2(ii)(iii) 得出，当 $X^\circ{}_+\neq\varnothing$ 时，$\forall\, x\in X$：$x\gg 0\Longleftrightarrow\langle X_+^*\backslash\{0\},x\rangle>0$. 以上结论给出将"向量不等式"转化为"标量不等式"的途径，在与向量序有关的问题中用处颇大.

**2.2.4 Krein 定理** 设 $A\subset X$ 是子空间，$K\subset X$ 是凸锥，$A\cap K^\circ\neq\varnothing$；$u$ 是 $A$ 上的连续线性泛函，$\langle u,A\cap K\rangle\geqslant 0$，则 $\exists\, v\in K^*$：$v|A=u$.

**证** 因可用如同证 Hahn-Banach 定理一样的延拓程序，不妨设 $X=A\oplus\mathrm{R}e$，$e\in A^c$. 取 $x_0\in A$，$\delta>0$，使 $\overline{B}_\delta(x_0)\subset K$，则 $\delta^{-1}|e|x_0\pm e\in K$，于是 $\pm\delta^{-1}|e|x_0=e\pm(\delta^{-1}|e|x_0\mp e)\in A\cap(e\pm K)\neq\varnothing$. $\forall\, x\in A\cap(e-K)$，$y\in A\cap(e+K)$，有 $y-x\in A\cap K$，从而 $u(y-x)\geqslant 0$. 这表明 $\langle u,A\cap(e-K)\rangle\leqslant\langle u,A\cap(e+K)\rangle$，于是 $\exists\,\xi\in\mathrm{R}$：$\langle u,A\cap(e-K)\rangle\leqslant\xi\leqslant\langle u,A\cap(e+K)\rangle$. 定义

$$v(a+\lambda e)=u(a)+\lambda\xi,\ a\in A,\ \lambda\in\mathrm{R},$$

则 $v$ 是 $X$ 上的线性泛函，$v|A=u$，由 $\xi$ 的取法易验知 $v(K)\geqslant 0$. 由 $\langle v,B_\delta(0)\rangle=\langle v,x_0-B_\delta(x_0)\rangle\leqslant v(x_0)$ 推出 $v$ 有界，因此 $v$ 如定理所求. □

给定一组锥 $K_i\subset X(1\leqslant i\leqslant n)$，确定 $(\bigcup K_i)^*$ 与 $(\bigcap K_i)^*$ 有其重要性. 直接看出 $(\bigcup K_i)^*=\bigcap K_i^*$，$(\bigcap K_i)^*\supset\mathrm{co}(\bigcup K_i^*)=\sum K_i^*$；要紧的是在什么条件下有等式

$$(\bigcap K_i)^* = \sum K_i^*. \tag{2}$$

**2.2.5 定理** 以下每个条件蕴涵等式(2):(i)$K_i(1\leqslant i\leqslant n)$是非空闭凸锥,$\sum K_i^*$ 弱$^*$ 闭;(ii)$K_i(1\leqslant i\leqslant n)$是开凸锥,$\bigcap K_i\neq\varnothing$;(iii)$n=2$,$K_1$ 是子空间,$K_2$ 为凸锥,且 $K_1\bigcap K_2^\circ\neq\varnothing$.

**证** 给定 $u\in(\bigcap K_i)^*$,只要证 $u\in\sum K_i^*$.

设条件(i)满足.若 $u\overline{\in}\sum K_i^*$,则由分离定理有 $x\in X$:$u(x)<\langle\sum K_i^*,x\rangle$,这推出 $u(x)<0$,而

$$x\in(\sum K_i^*)^* = (\bigcup K_i^*)^* = \bigcap K_i^{**} = \bigcap(X\bigcap K_i)$$

(2.2.3),从而 $x\in\bigcap K_i$,这与 $u\in(\bigcap K_i)^*$矛盾.

设条件(ii)满足.令 $D=\{(x,\cdots,x):x\in X\}\subset X^n$,$M=\Pi K_i$,$f(x,\cdots,x)=u(x)$,则 $D$ 是子空间,$M$ 是开凸锥,$D\bigcap M\neq\varnothing$,$f(D\bigcap M)\geqslant 0$.由 2.2.4,有 $g=(u_i)\in M^*$,$u_i\in X^*$,使得 $g|D=f$.这推出 $u=\sum u_i\in\sum K_i^*$.

设条件(iii)满足.对 $u|K_1$ 用 2.2.4 得 $v\in K_2^*$:$v|K_1=v|K_2$,于是 $u=(u-v)+v\in K_1^\perp+K_2^*$. □

现在考虑另一个类似的问题:设 $A\in L(X,Y)$,$K\subset Y$ 是一个锥,如何通过 $K^*$ 来表出$(A^{-1}K)^*$?直接由$\langle A^*K^*,A^{-1}K\rangle=\langle K^*,AA^{-1}K\rangle\geqslant 0$ 得出 $A^*K^*\subset(A^{-1}K)^*$.问题在于是否成立

$$(A^{-1}K)^* = A^*K^*. \tag{3}$$

**2.2.6 Farkas 引理** 设 $A\in L(X,Y)$,$K\subset Y$,则以下每个条件蕴涵等式(3):(i)$K$ 是闭凸锥,$A^*K^*\subset X^*$为弱$^*$闭;(ii)$K$ 是凸锥,$K^\circ\bigcap R(A)\neq\varnothing$.

**证** 取定 $u\in(A^{-1}K)^*$,只需证 $u\in A^*K^*$.若条件(i)满足而 $u\overline{\in}A^*K^*$,则由分离定理有 $x\in X$:$u(x)<\langle A^*K^*,x\rangle=\langle K^*,Ax\rangle$,这推出 $u(x)<0$,$Ax\in Y\bigcap K^{**}=K$,与 $u\in(A^{-1}K)^*$矛盾.若条件(ii)满足,则在 $X\times Y$ 中对子空间 Gr$A$ 与凸锥 $X\times K$ 应用 2.2.5 得

$$(\mathrm{Gr}A \cap (X \times K))^* = (\mathrm{Gr}A)^\perp + \{0\} \times K^*. \qquad (4)$$

易见$(u,0)\in(\mathrm{Gr}A\cap(X\times K))^*$，于是由(4)有$\lambda\in K^*$：$(u,0)=(u,-\lambda)+(0,\lambda)$，$(u,-\lambda)\in(\mathrm{Gr}A)^\perp$，这推出$u=A^*\lambda\in A^*K^*$. □

**注** 对$A\in L(\mathrm{R}^n,\mathrm{R}^m)$易直接看出$A^*\mathrm{R}_+^m$闭，于是由2.2.6有$(A^{-1}\mathrm{R}_+^m)^*=A^*\mathrm{R}_+^m$. 经典的Farkas引理原本指此.

Farkas引理是关于不等式系统可解性的一系列现代研究的开端，有许多重要应用. 在文献中，2.2.6的推广与变种颇多.

对偶锥概念的重要作用之一是通过$X_+$与$X_+^*$的联系相互推出彼此的性质，这种可能性基于两者的性质之间存在某种"对照表"，下面的结果给出这个对照表的一部分.

**2.2.7 定理** 设$X_+$为序锥. 则$X_+$是完全的$\Longleftrightarrow X_+^*$是序锥；$X_+$是再生的$\Longleftrightarrow X_+^*$是正规的；$X_+$是正规的$\Longleftrightarrow X_+^*$是再生的.

**证** 首先指出$X_+^*\neq\{0\}$（否则由2.2.3有$X_+=X\cap X_+^{**}=X$!），于是$X_+^*$是序锥$\Longleftrightarrow\{0\}=X_+^*\cap(-X_+^*)=X_+^\perp\Longleftrightarrow\overline{\mathrm{span}}X_+=X\Longleftrightarrow X_+$是完全的.

若$X_+$是再生的，则$\exists\,\delta>0$：$\delta B\subset C$，$B=B_1(0)$，$C$依2.1.3. $X_+^*$在$X^*$中导入序$\leqslant$；若$0\leqslant u\leqslant v$，则

$$\delta|u| = \sup u(\delta B) \leqslant \sup u(C) \leqslant \sup v(B) = |v|,$$

这表明$X_+^*$是正规的(2.1.4). 反之，若$X_+$不是再生的，则由2.1.3之证知$0\overline{\in}\bar{C}^\circ$，于是$\forall\, n\geqslant1$，$\exists\, x_n\in n^{-1}B\backslash\bar{C}$. 取$u_n\in X^*$：$u_n(\bar{C})<1<u_n(x_n)$，则$|u_n|>n$，且$u_n\in B^*+X_+^*$，此处$B^*=B_1(0)\subset X^*$（否则因$B^*+X_+^*$弱$^*$闭，由分离定理有$x\in X$：$\langle B^*+X_+^*,x\rangle<1<u_n(x)$，从而$x\in-X_+\cap B\subset C$，得出矛盾）. 同理$-u_n\in B^*+X_+^*$，这得出$X_+^*$非正规(2.1.4).

若$X_+^*$是再生的，则$X_+^{**}$是$X^{**}$中的正规锥，从而$X_+=X\cap X_+^{**}$正规. 反之，若$X_+^*$不是再生的，$D^*=B^*\cap X_+^*$，则$D^*-D^*$不是$X^*$的0-邻域（参考2.1.3之证）. 类似于上段证明可得出$x_n\in X$：$|x_n|\geqslant n$，$\langle D^*-D^*,x_n\rangle<1\,(n\geqslant1)$，这得出$\{x_n\}\subset(B+X_+)\cap$

$(B-X_+)$，从而 $X_+$ 非正规(2.1.4). □

## §3 正线性算子

设 $X$ 中已由序锥 $X_+$ 导入序 $\leqslant$. 显然 $L^+(X)\triangleq\{A\in L(X): AX_+\subset X_+\}$ 是 $L(X)$ 中的闭凸锥($X_+$ 完全 $\Rightarrow L^+(X)$ 为序锥)，它导入 $L(X)$ 中的序(亦记作 $\leqslant$). 称每个 $A\in L^+(X)$ 为正线性算子；若 $A(X_+\setminus\{0\})\subset X^\circ_+$，则说 $A$ 强正，记作 $A\gg 0$. 若 $X=\mathrm{R}^n$，则 $L^+(X)$ 是非负矩阵之集，$A=(a_{ij})\gg 0\Longleftrightarrow a_{ij}>0(1\leqslant i,j\leqslant n)$. 关于非负矩阵有以"Perron-Frobenius 理论"著称的丰富结果，其基本结论是(参考[21])：

**2.3.1 定理** 设 $A\in L^+(\mathrm{R}^n)$，$r=r_\sigma(A)$. 则 $r\in\sigma(A)$，且有 $x>0$：$Ax=rx$. 若 $A$ 不可约(即不存在置换矩阵 $P$，使 $P^{-1}AP=\begin{pmatrix}B & 0\\ C & D\end{pmatrix}$，$B,D$ 为方阵)，则 $r$ 是 $A$ 的简单特征值，且有 $x\gg 0$：$Ax=rx$；$\{\lambda\in\sigma(A):|\lambda|=r\}=\{r\exp(2k\pi i/m):1\leqslant k\leqslant m\}$，$m\leqslant n$. 若 $A\gg 0$，$r\neq\lambda\in\sigma(A)$，则必 $|\lambda|<r$. 若 $A\geqslant B\geqslant 0$，则 $r_\sigma(A)\geqslant r_\sigma(B)$.

下面的结果以"Krein-Rutman 定理"著称，它是 Perron-Frobenius 理论的无限维类似.

**2.3.2 定理** 设 $X_+$ 是完全的，$0\leqslant T\in CL(X)$，$r=r_\sigma(T)>0$. 则有 $0\neq x\in X_+$，$0\neq u\in X_+^*$：$Tx=rx$，$T^*u=ru$(称 $r$ 为 $T|X_+$ 的特征值)；当 $T\gg 0$ 时 $u$ 是严格正的.

**注** 因涉及谱论证，下面认定 $X$ 与 $T$ 皆已"复化". 约定 $T_\lambda=\lambda I-T$，$R_\lambda=T_\lambda^{-1}$.

**证** 首先设 $r\in\sigma(T)$，于是有 Laurent 展开

$$R_\lambda=\sum_{k=0}^{\infty}A_k(\lambda-r)^{k-n}=\sum_{k=0}^{\infty}\frac{T^k}{\lambda^{k+1}}\ (r<\lambda<r+\varepsilon),\quad (1)$$

其中 $n\geqslant 1$，$A_k\in L(X)$，$A_0\neq 0$，$\varepsilon>0$ 适当小. 由(1)推出

$$A_0=\lim_{\lambda\downarrow r}(\lambda-r)^nR_\lambda\geqslant 0,T_rA_0=A_0T_r=\lim_{\lambda\downarrow r}(\lambda-r)^nT_\lambda R_\lambda=0.\quad (2)$$

由 $X=\overline{\text{span}}X_+$ 与 $0\neq A_0\geqslant 0$ 有 $y>0:x\triangleq A_0y>0$；由(2)得 $T_rx=0$，即 $Tx=rx$. 必有 $v\in X_+^*:v(x)>0$(否则 $-x\in X_+$!). 令 $u=A_0^*v$，则 $T_r^*u=T_r^*A_0^*v=(A_0T_r)^*v=0$；$u\in X_+^*$；$u(y)=v(x)>0$. 若 $T\gg 0$ 而 $u$ 非严格正，则有 $z>0:u(z)=0$，于是对小的 $t>0$ 有 $0\leqslant u(Tz-ty)=-tu(y)<0$，得出矛盾.

今证 $r\in\sigma(T)$. 若有 $\lambda\in\sigma(T):|\lambda|=r,\lambda^m>0(m\geqslant 1)$，则对 $T^m$ 用已证结论得 $z>0:T^mz=\lambda^mz$，于是 $0<x\triangleq\sum_1^m r^{m-k}T^{k-1}z\in N(T_r)$，故 $r\in\sigma(T)$. 一般情况下取 $\lambda\in\sigma(T):|\lambda|=r$ 且 $\text{Re}\lambda$ 达最大，令 $\lambda_\varepsilon=\lambda+\varepsilon\lambda^2$，$T_\varepsilon=T+\varepsilon T^2$，则对小的 $\varepsilon>0$ 有 $|\lambda_\varepsilon|=r_\sigma(T_\varepsilon)$. 可选取任意小的 $\varepsilon>0$，使对某个 $m\geqslant 1$ 有 $\lambda_\varepsilon^m>0$；对这样的 $\lambda_\varepsilon$，由已证结论有 $x_\varepsilon>0:T_\varepsilon x_\varepsilon=|\lambda_\varepsilon|x_\varepsilon$，可设 $|x_\varepsilon|=1$. 由 $T$ 紧得出，有 $\varepsilon_n\downarrow 0$，使 $x_{\varepsilon_n}\to x>0$，于是 $Tx=|\lambda|x=rx$，$r\in\sigma(T)$. □

引入一个本章将多次使用的记号：

$$\alpha_x(y)=\sup\{\alpha\in\mathbf{R}_+:x+\alpha y\geqslant 0\},x\in X_+,y\in X. \quad (3)$$

容易验证：$x+\alpha_x(y)y\geqslant 0$；$x\gg 0\Rightarrow\alpha_x(y)>0$；$y\,\overline{\in}\,X_+\Rightarrow\alpha_x(y)<\infty$；$x+\alpha y\gg 0\Rightarrow\alpha<\alpha_x(y)$. 这些性质将被反复运用.

**2.3.3 定理** 设 $0\leqslant T\in CL(X)$，$r=r_\sigma(T)$. (i) $r>0$；$r$ 是 $T|X_+$ 的唯一特征值且为简单特征值. (ii) $r\neq\lambda\in\sigma(T)\Rightarrow|\lambda|<r$. (iii) 若 $y>0$，则方程 $T_\lambda x=y$ 有解 $x\geqslant 0\Leftrightarrow\lambda>r$；方程 $T_rx=\pm y$ 无解. (iv) 若 $S\in CL(X)$，$S\geqslant T[\forall\, x>0:Sx>Tx]$，则 $r_\sigma(S)\geqslant r[r_\sigma(S)>r]$.

**证** (i) 首先证 $r>0$. 取 $a\in X$，$\rho,\beta>0$，使 $B_\rho(a)\subset X_+$，$Ta\geqslant\beta a$. 由 $0\leqslant a-\rho T^na/|T^n||a|\leqslant a-\rho\beta^na/|T^n||a|$ 推出 $|T^n|\geqslant\rho\beta^n/|a|$，于是 $r\geqslant\beta>0$.

由 2.3.2，有 $e>0:Te=re\geqslant 0$. 若有 $\lambda\in\mathbf{R},x>0:Tx=\lambda x$，则 $0<\lambda\leqslant r$. 令 $\alpha=\alpha_x(-e)$，则 $x=\alpha e$(否则 $x-\alpha e>0$，从而 $0\ll T(x-\alpha e)=\lambda x-\alpha re$，这推出 $r<\lambda$!)，因而 $\lambda=r$，可见 $r$ 是 $T|X_+$ 的唯一实特征值. 若有 $\lambda\in\sigma(T)$，$x>0$，使 $Tx=\lambda x$，则用 2.3.2 的证法可得

$y>0:Ty=|\lambda|y$，于是$|\lambda|=r$. 由将在下面证明的结论(ii)，必$\lambda=r$，故$r$是$T|X_+$的唯一特征值.

证$N(T_r)=\mathrm{R}e$. 任给$x\in N(T_r)$. 若$x\geqslant 0$，则由上段所证知$x\in\mathrm{R}e$. 若$x\overline{\in}X_+$，则$\alpha\triangleq\alpha_e(x)<\infty$，$e=-\alpha x$，否则由$0\ll T(e+\alpha x)=r(e+\alpha x)$推出$\alpha<\alpha$!

证$N(T_r^2)=\mathrm{R}e$(如此则$r$是简单特征值). 任给$x\in N(T_r^2)$，有$T_rx=te, t\in\mathrm{R}$. 若$t>0$，则$x\geqslant 0$(否则$0<\beta\triangleq\alpha_e(x)<\infty$，$0\leqslant T(e+\beta x)=r(e+\beta x)-\beta te\ll r(e+\beta x)$，得出$\beta<\beta$!)，于是$rx=Tx+te\gg 0$，$0<\mu\triangleq\alpha_x(-e)<\infty$，$0\leqslant T(x-\mu e)=r(x-\mu e)-te\ll r(x-\mu e)$，这得出$\mu<\mu$! 因此$t\leqslant 0$；同理$t\geqslant 0$，于是$t=0$，$x\in N(T_r)=\mathrm{R}e$.

(ii)只就$r\neq\lambda\in\mathrm{R}\cap\sigma(T)$的情况证$|\lambda|<r$，完全的证明参看[202;p. 294]. 可设$\lambda\neq 0$，于是有$x\neq 0:T_\lambda x=0$. 因$\pm x\overline{\in}X_+$(否则由(i)有$\lambda=r$)，故$0<\alpha_\pm\triangleq\alpha_e(\pm x)<\infty$，$e$依(i)之证. 于是$0\ll T(e\pm\alpha_\pm x)=r(e\pm r^{-1}\lambda\alpha_\pm x)$. 若$\lambda>0$，则由$r^{-1}\lambda\alpha_+<\alpha_+$得$\lambda<r$. 若$\lambda<0$，则由$-r^{-1}\lambda\alpha_\pm<\alpha_\mp$得$\lambda>-r$，因此$|\lambda|<r$.

以下设$u$是依2.3.2的严格正的泛函且$T_r^*u=0$.

(iii)若有$x>0:T_\lambda x=y$，则$0<u(y)=u(T_\lambda x)=(\lambda-r)u(x)$，这推出$\lambda>r$. 反之，若$\lambda>r$，令$x=R_\lambda y$，则$T_\lambda x=y$，由(1)看出$x>0$. 因$\forall\ x\in X:u(T_rx)=0$，而$u(y)>0$，故方程$T_rx=\pm y$无解.

(iv)只需考虑$S\geqslant T$的情况. 令$\rho=r_\sigma(S)$，则由(i)有$x>0:Sx=\rho x$，于是$\rho u(x)=u(Sx)\geqslant u(Tx)=ru(x)$，这推出$\rho\geqslant r$. □

## §4 增映射

设$X$中已由序锥$X_+$导入序$\leqslant$. 给定$F:D\subset X\to X$，若$\forall\ x,y\in D:x\leqslant y\Rightarrow Fx\leqslant Fy$，则称$F$为增映射；若$x<y\Rightarrow Fx<Fy$[$Fx\ll Fy$]，则称$F$为严格增[强增]映射. 对于$F\in L(X)$，显然$F$是增映

射$\Longleftrightarrow F\geqslant 0$，$F$ 是强增映射$\Longleftrightarrow F\gg 0$（依§3）. 本节结果明显地受到单调实函数的某些直观性质的启发.

若 $D=[a,b]$，$a\leqslant b$，$F:D\to X$ 为增映射，则 $FD\subset D\Longleftrightarrow a\leqslant Fa$，$Fb\leqslant b$. 当 $a\leqslant Fa$，$Fb\leqslant b$ 时，迭代序列$\{F^n a\}$与$\{F^n b\}$分别为增序列与减序列；对 $F$ 或 $X_+$附加适当条件可得出 $F^n a$ 与 $F^n b$ 收敛于 $F$ 的不动点. 如此求解方程 $x=Fx$ 的方法称为“单调迭代方法”.

**2.4.1 定理** 设 $D=[a,b]\subset X$，$a\leqslant b$，$F:D\to D$ 为增映射，则以下每个条件推出 $F^n a\to x_*=\min \mathrm{Fix}F$，$F^n b\to x^*=\max \mathrm{Fix}F$：(i)$F\in\mathscr{K}$；(ii)$F\in\mathrm{SC}$，$X_+$正规；(iii)$F$ 连续，$X_+$正则.

**证** 只需考虑$x_n\triangleq F^n a$. 若 $x_n\to x_*$，则显然有 $x_*=\min\mathrm{Fix}F$，故只需证$\{x_n\}$收敛. 若 $F\in\mathscr{K}$，则$\{x_n\}$的每个子列有收敛子列；由单调性，$\{x_n\}$的任何收敛子列有同一极限 $x_*$，因此 $x_n\to x_*$. 若条件(ii)满足，则 $A\triangleq\{x_n:n\geqslant 0\}$有界，由 $\alpha(A)=\alpha(FA)\leqslant\alpha(F)\alpha(A)$ 推出 $\alpha(A)=0$，即 $A$ 相对紧，于是上面的论证仍然可用. 若 $X_+$正则，则直接得出$\{x_n\}$收敛. □

2.4.1 中的 $a$，$b$ 分别称为方程 $x=Fx$ 的下解与上解. 一般说来，求上、下解远比直接求解 $x=Fx$ 容易. 而一旦求得上、下解 $b$，$a$ 且 $b\geqslant a$，在一定条件下就可应用 2.4.1. 值得注意的是，2.4.1 不仅给出解的迭代算法，而且给出误差估计：$F^n a\leqslant x_*\leqslant x^*\leqslant F^n b$. 基于 2.4.1 的单调迭代方法被广泛应用于各种非线性方程的求解.

**2.4.2定义** 设$x_0\in D$，$T\in L(X)$. 若有$r>0$，使$x_0+X_+\cap B_r(0)\subset D$，当 $0\leqslant h\to 0$ 时 $\Delta F(x_0,h)-Th=o(|h|)$，则称 $T$ 为 $F$ 在 $x_0$ 沿 $X_+$的导数，记作 $F'_+(x_0)$. 若当 $x\in X_+$，$|x|\to\infty$时 $Fx$ 有定义且 $|Fx-Tx|=o(|x|)$，则称 $T$ 为 $F$ 在$\infty$沿 $X_+$的导数，记作 $F'_+(\infty)$. 类似地可定义 $F'_-(x_0)$与 $F'_-(\infty)$.

$F'_\pm(x_0)$显然是通常的“单侧导数”的推广. 若 $F'_+(x_0)$存在，则对任给 $h\in X_+$有

$$F'_+(x_0)h = \lim_{t \downarrow 0} \frac{1}{t}[F(x_0 + th) - Fx_0].$$

由此可见，若 $F$ 为增映射，或 $FD \subset Fx_0 + X_+$，则 $F'_+(x_0) \geqslant 0$. 类似地，若 $F'_+(\infty)$存在，$FD \subset X_+$，则 $F'_+(\infty) \geqslant 0$.

以下结果是 1.2.3 的一个类似.

**2.4.3 命题** 设 $T = F'_+(x_0)$ ($x_0 \in D$ 或 $x_0 = \infty$)存在，则 $\alpha(T|X_+) \leqslant \alpha(F)$；若 $X = \mathrm{span}X_+$，$F$ 全连续，则 $T \in CL(X)$.

**证** 只考虑 $x_0 \in D$ 的情况，$x_0 = \infty$ 的情况是类似的(参考 1.2.3 之证). $\forall\, \varepsilon > 0$，取 $\delta > 0$，使当 $h \in X_+$，$|h| \leqslant \delta$ 时 $|\Delta F(x_0, h) - Th| \leqslant \varepsilon |h|$. $\forall\, r > 0$，$A \subset X_+ \cap B_r(0)$，有

$$\alpha(TA) = \frac{r}{\delta}\alpha\left(T\left(\frac{\delta}{r}A\right)\right) \leqslant \frac{r}{\delta}\alpha\left(F\left(x_0 + \frac{\delta}{r}A\right) + \overline{B}_{\varepsilon\delta}(0)\right)$$

$$\leqslant \frac{r}{\delta}\left[\alpha(F)\alpha\left(\frac{\delta}{r}A\right) + 2\varepsilon\delta\right] = \alpha(F)\alpha(A) + 2r\varepsilon,$$

由此看出$\alpha(T|X_+) \leqslant \alpha(F)$. 若$X = \mathrm{span}X_+$，$\alpha(F) = 0$，则$\alpha(T|X_+) = 0$，$\exists\, \delta > 0$：$B_\delta(0) \subset C$，$C$依2.1.3. 显然$\alpha(TC) = 0$，于是$\alpha(TB_\delta(0)) = 0$，因此 $T \in CL(X)$. □

现在回到 2.4.1. 经常提出的问题是：是否存在 $x_0 \in \mathrm{Fix}F$，使 $x_* < x_0 < x^*$? 以$[x_*, x^*]$替代$[a, b]$，上述问题变成：若 $a, b \in \mathrm{Fix}F$，是否有 $x_0$ 使 $a < x_0 = Fx_0 < b$? 下面是一基本的结果.

**2.4.4 定理**(Amann，1976) 设 $D = [a, b] \subset X$，$F: D \to X$ 是紧增映射，$Fa = a \ll x < y \ll b = Fb$，$Fx \ll x$，$y \ll Fy$，$x, y$ 是给定的，则$\exists\, x_0$：$a < x_0 = Fx_0 < b$.

此定理的证明在 2.5.4 之后给出.

应用 2.4.4 的关键在于确定所要求的 $x, y$ 存在. 对 1 维特例的考虑启示出，$F$ 在 $a, b$ 两点的"单侧导数"可以起重要作用.

**2.4.5 命题** 设 $D = [a, b] \subset X$，$a \ll b$，$F: D \to X$ 为紧映射，$Fa = a$，$Fb = b$，$A = F'_+(a)$与$B = F'_-(b)$存在且$A, B \gg 0$，$r_a(A)$，$r_a(B) < 1$. 则$\exists\, x, y \in D$，使得

$$a \ll x < y \ll b, Fx \ll x, y \ll Fy. \tag{1}$$

**证** $A \gg 0$ 已暗含 $X^{\circ}_{+} \neq \varnothing$，因此 $X = \mathrm{span} X_{+}$ (2.1.3)，于是 $A, B \in CL(X)$ (2.4.3)，从而有 $e, e' \gg 0$：$Ae = r_{\sigma}(A)e$, $Be' = r_{\sigma}(B)e'$ (2.3.3). 取充分小的 $\bar{t} > 0$，令 $x = a + \bar{t}e$, $y = b - \bar{t}e'$，则 $a \ll x < y \ll b$. 由

$$F(a+te) = a + tAe + o(t) = a + tr_{\sigma}(A)e + o(t)$$
$$\ll a + te \quad (0 < t \to 0)$$

得 $Fx \ll x$；同理有 $y \ll Fy$，故(1)满足. □

为判定 $r_{\sigma}(A), r_{\sigma}(B) < 1$，可利用以下结果.

**2.4.6 命题** 设 $F: D \to D$ 为紧增映射，$D, x_{*}, x^{*}$ 依 2.4.1，$A = F'_{-}(x_{*})$ 与 $B = F'_{+}(x^{*})$ 存在且 $A, B \gg 0$，$a \ll x_{*} \leqslant x_{*} \ll b$，则 $r_{\sigma}(A), r_{\sigma}(B) \leqslant 1$；因此当 $r_{\sigma}(A), r_{\sigma}(B) \neq 1$ 时必 $r_{\sigma}(A), r_{\sigma}(B) < 1$.

**证** 如证 2.4.5 有 $e \gg 0$：$Ae = re$, $r = r_{\sigma}(A)$. 令 $x_t = x_{*} - te$. 若 $r > 1$，则当 $0 < t \to 0$ 时有 $a \leqslant x_t \ll x_{*}$，

$$a \leqslant Fa \leqslant Fx_t = x_{*} - tAe + o(t)$$
$$= x_{*} - tre + o(t) = x_t - t(r-1)e + o(t) \leqslant x_t.$$

于是由 2.4.1 有 $x_{*} \leqslant x_t \ll x_{*}$ $(t \downarrow 0)$，得出矛盾. 因此 $r_{\sigma}(A) \leqslant 1$；同理必 $r_{\sigma}(B) \leqslant 1$. □

以$\infty$代 $b$ 有 2.4.5 的以下类似：

**2.4.7 定理** 设 $D = a + X_{+}$，$F: D \to X$ 为全连续增映射，$Fa = a$，$A = F'_{+}(a)$ 与 $B = F'_{+}(\infty)$ 存在，且 $A, B \gg 0$，$r_{\sigma}(A) > 1$，$r_{\sigma}(B) < 1$；$X_{+}$ 正规. 则$\exists$ $x_0$：$a \ll x_0 = Fx_0$.

**证** 类似于 2.4.5 之证，有充分接近于 $a$ 的 $x$：$a \ll x \ll Fx$. 取 $e \gg 0$，使 $Be = re$, $r = r_{\sigma}(B)$. 令 $y_t = a + te$，则当 $t \to \infty$ 时 $y_t \geqslant x$，且 $y_t \geqslant 0$，

$$Fy_t = By_t + o(t) = Ba + tre + o(t)$$
$$= y_t + Ba - a + t(r-1)e + o(t) \leqslant y_t.$$

固定一个充分大的 $t$，令 $y = y_t$，则 $x \leqslant Fx \leqslant Fy \leqslant y$. 因 $X_{+}$ 正规，故 $[x, y]$ 有界. 在 $[x, y]$ 上应用 2.4.1 得出 $x_0 \in [x, y] \cap \mathrm{Fix} F$，于是 $a \ll x_0 = Fx_0$. □

## §5 不动点指数

上章所给出的度理论有一明显的缺陷，即所考虑的映射 $F$ 必须定义于某个开集 $\Omega\subset X$ 上，这就大大限制了度的应用. 若将度理论推广到定义于一定相对开集上的映射，应用范围就会大得多. 而这正是本节所要作的.

取定非空闭凸集 $K\subset X$. 任给相对开集 $\Omega\subset K$，约定以 $\partial\Omega$ 记 $\Omega$ 在 $K$ 中的边界. 由 1.2.5，$\mathrm{id}:K\to K$ 有扩张 $P\in C(X,X)$，使 $PX=K$，即 $P$ 是从 $X$ 到 $K$ 的收缩. 任给紧映射 $F:\overline{\Omega}\to K$，$FP:\overline{P^{-1}\Omega}\to X$ 是 $F$ 的紧扩张，对 $I-FP$ 可以定义度. 这引出度概念的以下推广：

**2.5.1 定义** 任给相对开集 $\Omega\subset K$ 与紧映射 $F:\overline{\Omega}\to K$，$\partial\Omega\cap\mathrm{Fix}F=\varnothing$，取收缩 $P:X\to K$ 与充分大的 $r>0$，称

$$i(F,\Omega)\triangleq d(I-FP,B_r(0)\cap P^{-1}\Omega,0)\tag{1}$$

为 $F$ 关于 $K$ 的不动点指数，也写作 $i_K(F,\Omega)$.

**注** 1° 不难看出(1)之右端有定义，且由$(D_3)(D_4)$(依§1.3，下同)推出它与 $P,r$ 之选取无关.

2° 当 $\Omega$ 有界时可进而定义集压缩映射 $F:\overline{\Omega}\to K$ 的不动点指数，且保持下面给出的大多数结论. 但为简便起见限于考虑紧映射.

3° 若 $K=X,P=I,\Omega$ 有界，则 $i(F,\Omega)=d(I-F,\Omega,0)$. 可见不动点指数包含度为其特例.

本节的大部分内容可与§1.3对照起来理解. 首先，对应于度性质$(D_1)\sim(D_7)$，下面给出不动点指数的性质$(I_1)\sim(I_6)$.

**2.5.2 定理** 设 $\Omega\subset K$ 为相对开集，$F,G:\overline{\Omega}\to K$ 为紧映射，$\partial\Omega\cap\mathrm{Fix}F=\varnothing$，则以下结论成立：$(I_1)$若 $Fx\equiv y\in\Omega$，则 $i(F,\Omega)=1$；$(I_2)$若 $\Omega_j(1\leqslant j\leqslant n)$是 $\Omega$ 的互不相交开子集，$(\overline{\Omega}\backslash\bigcup\Omega_j)\cap\mathrm{Fix}F=$

$\varnothing$，则 $i(F,\Omega)=\sum i(F,\Omega_j)$；($I_3$)若 $H:J\times\overline{\Omega}\to K$ 是紧同伦，$\partial\Omega\cap\mathrm{Fix}H_t=\varnothing(\forall\ t\in J)$，则 $i(H_t,\Omega)$ 与 $t$ 无关；($I_4$)若 $A\subset\overline{\Omega}$ 闭，$(A\cup\partial\Omega)\cap\mathrm{Fix}F=\varnothing$，则 $i(F,\Omega)=i(F,\Omega\backslash A)$；$i(F,\Omega)\neq 0\Rightarrow\mathrm{Fix}F\neq\varnothing$；($I_5$)若 $\forall\ x\in\partial\Omega$：$x\overline{\in}[Fx,Gx]$（特别，若 $\forall\ x\in\partial\Omega$：$|Fx-Gx|<|x-Fx|$ 或 $F|\partial\Omega=G|\partial\Omega$），则 $i(F,\Omega)=i(G,\Omega)$；($I_6$)归约性质：若 $F\overline{\Omega}\subset M\subset K$，$M$ 为闭凸集，则 $i_K(F,\Omega)=i_M(F,M\cap\Omega)$.

性质($I_1$)～($I_5$)明显地对应于度性质($D_1$)～($D_5$). ($I_1$)～($I_3$)分别从($D_1$)～($D_3$)推出. 例如考虑($I_3$)：令 $\tilde{H}(t,x)=H(t,Px)$，$P$ 依 2.5.1，则 $\tilde{H}:J\times\overline{B_r(0)\times P^{-1}\Omega}\to X$ 是紧同伦，于是由($D_3$)推出

$$\begin{aligned}i(H_t,\Omega)&=d(I-H_tP,B_r(0)\cap P^{-1}\Omega,0)\\&=d(I-\tilde{H}_t,B_r(0)\cap P^{-1}\Omega,0)\end{aligned}$$

与 $t$ 无关. 如同从($D_1$)～($D_3$)推出($D_4$)($D_5$)，可从($I_1$)～($I_3$)推出($I_4$)($I_5$). ($I_6$)亦可用类似的方法得出.

**2.5.3 推论** 若 $F:K\to K$ 为紧映射，则 $i_K(F,K)=1$.

**证** 任取 $x_0\in K$，令 $H(t,x)=tFx+t'x_0$，则 $H:J\times K\to K$ 为紧同伦，于是由($I_1$)($I_3$)有 $i_K(F,K)=i_K(H_0,K)=1$. □

**注** 2.5.3 蕴涵 Schauder 不动点定理(1.4.3). 当 $K$ 有界时 2.5.3 亦适用于集压缩映射 $F$，因此可以说 2.5.3 亦蕴涵 Darbo 不动点定理. 由此可见，从不动点指数得出不动点定理是更为直接的.

**2.5.4 三不动点定理** 设 $F:K\to K$ 为紧映射，$D_1,D_2$ 是 $K$ 的互不相交非空闭凸子集，$FD_j\subset D_j$，$\Omega_j$ 记 $D_j$ 在 $K$ 中的内部，$(D_j\backslash\Omega_j)\cap\mathrm{Fix}F=\varnothing$，$j=1,2$，则 $F$ 在 $\Omega_1,\Omega_2$ 与 $K\backslash(D_1\cup D_2)$ 中都有不动点.

**证** 由 Schauder 不动点定理知 $F$ 在 $\Omega_1,\Omega_2$ 中有不动点. 用($I_2$)($I_4$)($I_6$)及 2.5.3 有

$$\begin{aligned}i_K(F,K\backslash(D_1\cup D_2))&=i_K(F,K)-\sum_j i_K(F,\Omega_j)\\&=1-\sum_j i_{D_j}(F,\Omega_j)=1-\sum_j i_{D_j}(F,D_j)\end{aligned}$$

$= 1 - 2 \neq 0$,

可见 $F$ 在 $K\backslash(D_1 \cup D_2)$ 中亦有不动点. □

**注** 设 $x, y$ 如 2.4.4,对 $K=[a,b]$, $D_1=[a,x]$ 与 $D_2=[y,b]$ 应用 2.5.4 得出所要不动点 $x_0$ 存在.

以下设 $K$ 是序锥且在 $X$ 中导入序 $\leqslant$, $F:D\to K$ 为紧映射, $0\in\Omega\subset\bar{\Omega}\subset D\subset K$, $\Omega$ 是 $K$ 的相对开子集;约定 $K_r=B_r(0)\cap K$. 下面两个定理的形式与作用恰与 1.3.4, 1.3.5 相当.

**2.5.5 定理** 以下每个条件蕴涵 $i(F,\Omega)=1$:

$(H_1)$ Leray-Schauder 条件(LS): $(\lambda,x)\in J\times\partial\Omega\Rightarrow x\neq\lambda Fx$;

$(H_2^*)$ $\Omega=K_r$, $r>0$ 充分小, $F(0)=0$, $A=F'_+(0)$ 存在且 $A|K$ 无 $\geqslant 1$ 的实特征值;

$(H_3)$ $D=K$, $\Omega=K_\rho$, $\rho$ 充分大, $B=F'_+(\infty)$ 存在且 $B|K$ 无 $\geqslant 1$ 的实特征值;

$(H_4)$ $\forall\ x\in\partial\Omega$: $|Fx|\leqslant|x|$ 且 $Fx\neq x$;

$(H_5)$ $\forall\ x\in\partial\Omega$: $x\not\leqslant Fx$ 或 $\forall\ x\in\partial\Omega$: $Fx<x$.

**证** 若 $(H_1)$ 满足,则由 $(I_1)(I_3)$ 有 $i(F,\Omega)=i(\lambda F,\Omega)=i(0,\Omega)=1(\lambda\in J)$. 若 $(H_2)$ 满足,则 $\alpha(A|K)=0$(2.4.3),因此 $(I-A)\cdot(\partial K_1)$ 闭(1.2.4),于是 $2\varepsilon\triangleq\inf\{|x-Ax|: x\in\partial K_1\}>0$. 可设 $\forall\ x\in\partial K_r$: $|Fx-Ax|\leqslant\varepsilon|x|=\varepsilon r$, 于是 $|Fx-Ax|<2\varepsilon r\leqslant|x-Ax|$ $(\forall\ x\in\partial K_r)$. 用 $(I_5)$ 并注意 $A$ 满足 $(H_1)$ 得 $i(F,K_r)=i(A,K_r)=1$. 类似地可证 $(H_3)\Rightarrow i(F,K_\rho)=1$. 显然 $(H_4)\Rightarrow(H_1)$, $(H_5)\Rightarrow(H_1)$, 于是定理得证. □

**2.5.6 定理** 设 $\Omega$ 有界,则以下每个条件蕴涵 $i(F,\Omega)=0$:

$(H'_1)$ $\exists\ e>0$: $(I-F)(\partial\Omega)\cap \mathrm{R}_+e=\varnothing$;

$(H'_2)$ $\Omega=K_r$, $r>0$ 充分小, $F(0)=0$, $A=F'_+(0)$ 存在且 $A|K$ 有 $>1$ 的特征值 $\lambda$ 而无特征值 1;

$(H'_3)$ $D=K$, $\Omega=K_\rho$, $\rho$ 充分大, $B=F'_+(\infty)$ 存在且 $B|K$ 有特征值 $\lambda>1$ 而无特征值 1;

$(H'_4)$ $\forall\ x\in\partial\Omega$: $|x|\leqslant|Fx|$ 且 $Fx\neq x$;

$(H'_5)$ $\forall\ x\in\partial\Omega: Fx\not\leqslant x$，或$\forall\ x\in\partial\Omega: x<Fx$；

$(H'_6)$ $d(0,F(\partial\Omega))>0, \forall\ \lambda\geqslant 1: \partial\Omega\cap \mathrm{Fix}(\lambda F)=\varnothing$.

**证** 用 1.3.5 的证法可证$(H'_1)\Rightarrow i(F,\Omega)=0$.

若$(H'_2)$满足，则$i(F,K_r)=i(A,K_r)$（参考 2.5.5 之证）. 取$e>0$：$Ae=\lambda e$，则必$(I-A)(\partial K_r)\cap \mathrm{R}_+e=\varnothing$（从而$i(A,K_r)=0$），否则$\exists\ t\geqslant 0, x\in\partial K_r: x-Ax=te$. 设$\alpha\triangleq\alpha_x(-e)$依 § 3(3)，则$0\leqslant A(x-\alpha e)=x-(t+\alpha\lambda)e$，这推出$t+\alpha\lambda\leqslant\alpha$，从而$t=\alpha=0, x=Ax$，这与 1 不是$A|K$的特征值矛盾. 因此$i(F,K_r)=0$. 类似地可证$(H'_3)\Rightarrow i(F,K_\rho)=0$.

易验明$(H'_5)\Rightarrow(H'_1)$，$(H'_4)\Rightarrow(H'_6)$. 今设$(H'_6)$满足. 取$F|\partial\Omega$的紧扩张$G$，使$G\overline{\Omega}\subset\overline{\mathrm{co}}F(\partial\Omega)\triangleq M$. 由$M$紧推出$Y\triangleq\overline{\mathrm{span}}M$可分，于是由 2.2.2 有$u\in Y^*: u(K_0\backslash\{0\})>0, K_0=K\cap Y$. 因$C\triangleq\overline{F(\partial\Omega)}$紧且$0\overline{\in}C\subset K_0$，故$u(C)\geqslant\varepsilon>0$，于是$u(M)\geqslant\varepsilon$，从而$0\overline{\in}M$. 这推出对充分大的$t$有$\overline{\Omega}\cap\mathrm{Fix}(F+tG)=\varnothing$，于是$i(F,\Omega)=i(F+tG,\Omega)=0$. □

注意条件$(H'_1)\sim(H'_6)$可与 2.5.5 的条件$(H_1)\sim(H_6)$对照，而条件$(H_1)\sim(H_4)$则可与 1.3.4 的条件$(H_1)\sim(H_4)$类比；$(H'_1)(H'_4)(H'_6)$亦可与 1.3.5 中的$(H'_1)(H'_4)(H'_3)$类比.

适当组合 2.5.5 与 2.5.6 中的条件，可得出一系列不动点结果，仅举一典型例子如下：

**2.5.7 推论** 设$0<r<\rho$，$F:\overline{K}_\rho\to K$为紧映射. 若$\forall\ x\in\partial K_r, y\in\partial K_\rho: |Fx|\leqslant|x|, |y|\leqslant|Fy|$（或$|x|\leqslant|Fx|, |Fy|\leqslant|y|$），则$F$有不动点$x\in\overline{K}_\rho\backslash K_r$.

通称 2.5.7 为"锥拉伸（或锥压缩）不动点定理".

注意 2.5.5 中条件$(H_1)$无需假定$K$为锥，故有

**2.5.8 定理** 设$K\subset X$为闭凸集，$\Omega\subset K$为相对开集，$0\in\Omega$，$F:\overline{\Omega}\to K$为紧映射，$\forall\ (\lambda,x)\in(0,1)\times\partial\Omega: x\neq\lambda Fx$，则$\mathrm{Fix}F\neq\varnothing$.

对于$F$为紧映射的情况，2.5.8 蕴涵了 1.4.1.

## §6 某些应用

本章的结果已广泛应用于各种非线性方程的正解问题(如见[2,3,27,53,58,67,130,159]),应用的方式与技巧是如此多样,要作简单的概括是不可能的.本节考虑的几个例子只能起说明作用,未必具有代表性.

首先考虑在上章中讨论过的 BVP:

$$\begin{cases} Lx \triangleq x^{(n)} - \sum A_i(t)x^{(i-1)} = f(t,\widetilde{x})(t \in J); & (1) \\ Bx = 0. & (2) \end{cases}$$

沿用§1.6中的所有记号,且设§1.6中的基本假设$(H_1)\sim(H_3)$满足,于是 BVP(1)(2)转化为不动点方程 $x=Tx, T=KF: X_0 \to X_0$ 全连续.令 $X_+ = X_0 \cap C(J, R^m_+)$,则 $X_+$ 是 $X_0$ 中的序锥,它所定义的序$\leqslant$具有自然的含义.若 $x=Tx \in X_+$,则称 $x$ 为 BVP(1)(2)的非负解,当 $x \neq 0$ 时称 $x$ 为正解.若 $TX_+ \subset X_+$,则可利用上节导出的种种不动点结论来得出(1)(2)非负解的存在性.例如,由2.5.8得出对应于1.6.4的以下结果:

**2.6.1 定理** 设 $TX_+ \subset X_+$,存在 $X_+$ 的有界相对开子集 $\Omega$,使得 $0 \in \Omega, \forall \lambda \in (0,1)$:BVP

$$Lx = \lambda f(t,\widetilde{x}) \ (t \in J),\ Bx = 0 \qquad (3)_\lambda$$

的任何正解 $x \overline{\in} \partial\Omega$,$\partial\Omega$ 记 $\Omega$ 在 $X_+$ 中的边界,则 BVP(1)(2)有非负解.

以下推论恰与1.6.5相当.

**2.6.2 推论** 若 $TX_+ \subset X_+$,且

$$\sup\{|x|_X : x \in X_+, \exists \lambda \in (0,1), x \text{ 满足} (3)_\lambda\} < \infty, \qquad (4)$$

则 BVP(1)(2)有非负解.

若能求得 $T$ 的明显表达式:

$$Tx(t) = \int_0^1 G(t,s)f(s,\widetilde{x}(s))\mathrm{d}s, \qquad (5)$$

其中 $G(t,s)$是 Green 函数，则可利用(5)来判定是否有 $TX_+\subset X_+$. 不过，(5)式未必容易得出；即使可写出亦可能因其过繁而难于利用. 因此，只要有可能，就直接从已给资料判定 $TX_+\subset X_+$是可取的. 下面以一简单例子作说明.

**2.6.3 例** 考虑 2 阶 BVP

$$x''=f(t,x,x')(t\in J),\ x(0)=0,x(\xi)=x(1), \tag{6}$$

其中 $\xi\in(0,1)$是给定的. 设 $f:J\times R^{2m}\to R^m$ 满足 Caratheodory 条件以及条件

($H_4$) $\exists\ p,r\in L^1(J,R_+),q\in L^2(J,R_+),\alpha>0$，使得

$$|f(t,x,y)|\leqslant p(t)|x|+q(t)|y|+r(t),$$
$$(t,x,y)\in J\times R_+^m\times R^m; \tag{7}$$

$$|p|_1+|q|_2+\frac{2\alpha^2}{3\pi}\sqrt{3(1-\xi)}<1. \tag{8}$$

$$f(t,x,y)\leqslant\alpha^2x,(t,x,y)\in J\times R_+^m\times R^m. \tag{9}$$

令 $X=C^1(J,R^m)$, $|x|_X=\max\{|x|_0,|x'|_0\}$, $X_0=\{x\in X:x(0)=0,x(\xi)=x(1)\}$, $X_+=X_0\cap C(J,R_+^m)$, $Lx=x''-\alpha^2x$, $Fx(t)=f(t,x,x')-\alpha^2x(t)$, $T=L^{-1}F:X_0\to X_0$. 经某些计算后得出

$$Tx(t)=\int_0^1G(t,s)[f(s,x(s),x'(s))-\alpha^2x(s)]ds, \tag{10}$$

其中 $G(t,s)$决定于下式：

$$\alpha(\mathrm{sh}\alpha\xi-\mathrm{sh}\alpha)G(t,s)$$
$$=\begin{cases}2\mathrm{sh}\,\dfrac{\alpha(1-\xi)}{2}\mathrm{sh}\alpha s\mathrm{ch}\,\dfrac{\alpha(1+\xi-2t)}{2}, & s\leqslant t\leqslant\xi;\\ 2\mathrm{sh}\,\dfrac{\alpha(1-\xi)}{2}\mathrm{sh}\alpha t\mathrm{ch}\,\dfrac{\alpha(1+\xi-2s)}{2}, & t<s\leqslant\xi;\\ \mathrm{sh}\alpha t\mathrm{sh}\alpha(1-s)-\mathrm{sh}\alpha(t-s)(\mathrm{sh}\alpha-\mathrm{sh}\alpha\xi), & \xi\leqslant s<t;\\ \mathrm{sh}\alpha t\mathrm{sh}\alpha(1-s), & \xi\leqslant t\leqslant s.\end{cases} \tag{11}$$

由(11)可验知 $G(t,s)\leqslant0(t,s\in J)$，这结合(9)得出 $TX_+\subset X_+$. 设 $x\in X_+$满足

$$x''-\alpha^2x=\lambda[f(t,x,x')-\alpha^2x]\ (t\in J), \tag{12}$$

其中 $\lambda\in J$，今仿照 1.6.7 估计 $|x|_X$. 令 $\mu=|x'|_2$，则

$$\mu^2=(x(t),x'(t))|_0^1-\int_0^1(x(t),x''(t))\mathrm{d}t$$
$$=(x(1),x'(1))-(1-\lambda)\alpha^2|x|_2^2-\lambda\int_0^1(x(t),f(t,x(t),$$
$$x'(t))\mathrm{d}t$$
$$\leqslant(x(1),x'(1))+\int_0^1|x(t)||f(t,x(t),x'(t))|\mathrm{d}t$$
$$\triangleq A_1+A_2. \tag{13}$$

显然 $|x|_0\leqslant\mu$. 由 Wirtiager 不等式([63;7.7])有 $|x|_2\leqslant(2/\pi)\mu$. 由

$$0=\int_\xi^1 x'(t)\mathrm{d}t=\int_\xi^1[x'(1)-\int_t^1 x''(s)\mathrm{d}s]\mathrm{d}t$$

得
$$x'(1)=\frac{1}{1-\xi}\int_\xi^1\mathrm{d}t\int_t^1 x''(s)\mathrm{d}s.$$

由 $x''-\alpha^2x=\lambda Fx\leqslant 0$ 有 $x''\leqslant\alpha^2x$. 于是

$$A_1=(x(1),x'(1))=\frac{1}{1-\xi}\int_\xi^1\mathrm{d}t\int_t^1(x(1),x''(s))\mathrm{d}s$$
$$\leqslant\frac{\alpha^2}{1-\xi}\int_\xi^1\mathrm{d}t\int_t^1(x(1),x(s))\mathrm{d}s\leqslant\frac{\alpha^2\mu}{1-\xi}\int_\xi^1\mathrm{d}t\int_t^1|x(s)|\mathrm{d}s$$
$$=\frac{\alpha^2\mu}{1-\xi}\int_\xi^1(t-\xi)|x(t)|\mathrm{d}t$$
$$\leqslant\frac{\alpha^2\mu}{1-\xi}\sqrt{\frac{(1-\xi)^3}{3}}|x|_2\leqslant\frac{2\alpha^2\mu^2}{3\pi}\sqrt{3(1-\xi)}. \tag{14}$$

其次用(7)估计 $A_2$:

$$A_2\leqslant\int_0^1[p(t)|x(t)|^2+q(t)|x(t)||x'(t)|+r(t)|x(t)|]\mathrm{d}t$$
$$\leqslant|p|_1|x|_0^2+|q|_2|x'|_2|x|_0+|r|_1|x|_0$$
$$\leqslant\mu^2(|p|_1+|q|_2)+\mu|r|_1. \tag{15}$$

综合(13)(14)(15)得(注意用(8))

$$\mu\leqslant\left(1-|p|_1-|q|_2-\frac{2\alpha^2}{3\pi}\sqrt{3(1-\xi)}\right)^{-1}|r|_1. \tag{16}$$

结合(7)与(12)有

$$|x''|_1 \leqslant \alpha^2 |x|_0 + |p|_1 |x|_0 + |q|_2 |x'|_2 + |r|_1$$
$$\leqslant \mu(\alpha^2 + |p|_1 + |q|_2) + |r|_1. \tag{17}$$

设 $x=(x_i)(1\leqslant i\leqslant m)$. 由 $x(\xi)=x(1)$ 有 $\tau_i\in(\xi,1)$: $x'_i(\tau_i)=0$, 由此推出 $|x'_i(t)|\leqslant|x''|_1$, 从而 $|x'|_0\leqslant m|x''|_1$. 这与(16)(17)及 $|x|_0\leqslant\mu$ 一起得出 $|x|_X$ 不超过某个与 $\lambda,x$ 无关的常数, 因此由 2.6.2 得出 BVP(6)有非负解.

在上面讨论中, 除验证 $TX_+\subset X_+$ 之外, 并不用到 $G(t,s)$. 实际上, 不用 $G(t,s)$ 亦可验明 $TX_+\subset X_+$. 任给 $x\in X_+$, 令 $z=Tx$, 则 $z$ 满足

$$z'' - \alpha^2 z = Fx, z(0) = 0, z(\xi) = z(1). \tag{18}$$

由 $Fx\leqslant 0$ 得 $z''\leqslant\alpha^2 z$. 若 $z\overline{\in}X_+$, $z=(z_i)$, 则有某个 $z_i\ngeqslant 0$, 于是有极大子区间 $\Delta=(\sigma,\tau)\subset J$, 在 $\Delta$ 内 $z_i(t)<0$. 必定 $z_i(\sigma)=0$, $z'_i(\sigma)\leqslant 0$. 因在 $\Delta$ 内 $z''_i(t)\leqslant\alpha^2 z_i(t)<0$, 故 $z'_i(t)$ 与 $z_i(t)$ 在 $\Delta$ 内严格下降, $z_i(\tau)<0$, 这推出 $\tau=1$. 但这与 $z_i(\xi)=z_i(1)$ 矛盾.

下面考虑 Hammerstein 型积分方程

$$x(t) = \int_0^1 G(t,s)f(s,x(s))\mathrm{d}s \triangleq Tx(t), \tag{19}$$

其中 $G: J\times J\to\mathrm{R}_+$, $f: J\times\mathrm{R}_+\to\mathrm{R}_+$. 假定 $G$ 与 $f$ 的选取使 $T$: $C(J,\mathrm{R}_+)\subset X\to X=C(J,\mathrm{R})$ 为全连续映射(为此需加于 $G,f$ 的条件是很宽的). 设

$$u(t) = \inf_{s,\tau\in J} G(t,s)/G(\tau,s) \quad (t\in J) \tag{20}$$

有定义且 $u(t)\not\equiv 0$. 令 $X_+=\{x\in X: x(t)\geqslant u(t)|x|_0 (\forall t\in J)\}$, 则可验证 $X_+$ 是 $X$ 中的序锥. 与看来更自然的锥 $C(J,\mathrm{R}_+)$ 相比, $X_+$ 有某些优点. 任给 $x\in X_+$, 因 $Tx\in C(J,\mathrm{R}_+)$, 故有 $\tau\in J$: $|Tx|_0=Tx(\tau)$. 于是

$$Tx(t) \geqslant \int_0^1 u(t)G(\tau,s)f(s,x(s))\mathrm{d}s = u(t)|Tx|_0,$$

可见 $Tx\in X_+$. 故有 $TX_+\subset X_+$. 为应用上节的结果，作以下假设*：

($H_5$) $\exists\ \rho_1>0, 0<a<b<1$，增函数 $\varphi_1:[0,\rho_1]\to\mathbf{R}$，可测函数 $q_1:J\to\mathbf{R}_+$，使得 $m\triangleq\inf\limits_{a\leqslant t\leqslant b}u(t)>0$，

$$f(t,x)\geqslant q_1(t)\varphi_1(x),(t,x)\in[a,b]\times[m\rho_1,\rho_1] \tag{21}$$

$$\exists\ t\in J:\varphi_1(m\rho_1)\int_a^b G(t,s)q_1(s)\mathrm{d}s\geqslant\rho_1. \tag{22}$$

($H_6$) $\exists\ \rho_2>0$，增函数 $\varphi_2:[0,\rho_2]\to\mathbf{R}$ 与可测函数 $q_2:J\to\mathbf{R}_+$，使得

$$f(t,x)\leqslant q_2(t)\varphi_2(x),(t,x)\in J\times[0,\rho_2]; \tag{23}$$

$$\varphi_2(\rho_2)\int_0^1 G(t,s)q_2(s)\mathrm{d}s\leqslant\rho_2(\forall\ t\in J). \tag{24}$$

令 $\Omega_i=\{x\in X_+:|x|_0<\rho_i\}$，则 $\Omega_i$ 是 $X_+$ 中的相对开集. 若 $x\in\partial\Omega_1$，则 $|x|_0=\rho_1$，$x(t)\geqslant\rho_1u(t)$ ($\forall\ t\in J$)，当 $a\leqslant t\leqslant b$ 时 $m\rho_1\leqslant x(t)\leqslant\rho_1$. 取定 $t\in J$ 使(22)成立，则

$$\begin{aligned}|Tx|_0&\geqslant\int_0^1 G(t,s)f(s,x(s))\mathrm{d}s\\&\geqslant\int_a^b G(t,s)f(s,x(s))\mathrm{d}s\\&\geqslant\int_a^b G(t,s)q_1(s)\varphi(x(s))\mathrm{d}s\\&\geqslant\varphi(m\rho_1)\int_a^b G(t,s)q_1(s)\mathrm{d}s\geqslant\rho_1,\end{aligned}$$

可见 $|Tx|_0\geqslant|x|_0$. 若 $x\in\partial\Omega_2$，则 $\rho_2u(t)\leqslant x(t)\leqslant\rho_2$，于是由(23)(24)有

$$\begin{aligned}Tx(t)&\leqslant\int_0^1 G(t,s)q_2(s)\varphi_2(x(s))\mathrm{d}s\\&\leqslant\varphi_2(\rho_2)\int_0^1 G(t,s)q_2(s)\mathrm{d}s\leqslant\rho_2,\end{aligned}$$

---

* ($H_5$)($H_6$)的某种类似形式是王春林告知作者的.

这得出 $\|Tx\|_0 \leqslant \|x\|_0$. 由此可见，当 $\rho_1 < \rho_2$ 时 $T$（相对于 $\Omega_2 \backslash \bar{\Omega}_1$）为"锥压缩"；当 $\rho_1 > \rho_2$ 时 $T$（相对于 $\Omega_1 \backslash \bar{\Omega}_2$）为"锥拉伸"，于是可应用 2.5.7 得出：

**2.6.4 定理** 若条件($H_5$)($H_6$)满足且 $\rho_1 \neq \rho_2$，则方程(19)有正解.

**2.6.5 例** 考虑 BVP

$$x'' + q(t)f(x) \ (t \in J), \ x(0) = x(1) = 0. \tag{25}$$

熟知 BVP(25)可化为积分方程

$$x(t) = \int_0^1 G(t,s)q(s)f(x(s))\mathrm{d}s, \tag{26}$$

其中

$$G(t,s) = \begin{cases} (1-t)s, s \leqslant t; \\ (1-s)t, s > t. \end{cases}$$

依(20)算得 $u(t) = 2^{-1} - |t - 2^{-1}|$. 假定 $f \in C(\mathrm{R}_+, \mathrm{R}_+)$，$q: J \to \mathrm{R}_+$ 可测，且 $\int_0^1 q(t)\mathrm{d}t > 0$，$\int_0^1 t(1-t)q(t)\mathrm{d}t < \infty$. 令

$$\alpha_1 = \overline{\lim_{x \downarrow 0}} \frac{f(x)}{x}, \alpha_2 = \underline{\lim_{x \downarrow 0}} \frac{f(x)}{x},$$

$$\beta_1 = \underline{\lim_{x \to \infty}} \frac{f(x)}{x}, \beta_2 = \overline{\lim_{x \to \infty}} \frac{f(x)}{x}.$$

设定条件：

($H_7$) 存在 $[a,b] \subset (0,1)$，$m = \min\{u(a), u(b)\}$，使得

$$\alpha_1 \int_0^1 (1-t)tq(t)\mathrm{d}t < 1 < m\beta_1 \int_a^b G\left(\frac{1}{2}, t\right) q(t)\mathrm{d}t. \tag{27}$$

($H_8$) 存在 $[a,b] \subset (0,1)$，$m = \min\{u(a), u(b)\}$，使得

$$\beta_2 \int_0^1 (1-t)tq(t)\mathrm{d}t < 1 < m\alpha_2 \int_a^b G\left(\frac{1}{2}, t\right) q(t)\mathrm{d}t. \tag{28}$$

若($H_7$)满足，则必 $\alpha_1 < \infty$，$\beta_1 > 0$，且∃ $\alpha'_1 \in (\alpha_1, \infty)$，$\beta'_1 \in (0, \beta_1)$：

$$\alpha'_1 \int_0^1 (1-t)tq(t)\mathrm{d}t < 1 < m\beta'_1 \int_a^b G\left(\frac{1}{2}, t\right) q(t)\mathrm{d}t. \tag{29}$$

取 $\rho_1>0$，使得 $\forall\ x\geqslant m\rho_1: f(x)\geqslant\beta'_1 x\triangleq\varphi_1(x)$；取 $\rho_2\in(0,\rho_1)$，使得 $\forall\ x\in[0,\rho_2]: f(x)\leqslant\alpha'_1 x\triangleq\varphi_2(x)$. 利用(29)及明显的事实 $G(t,s)\leqslant(1-s)s$ 容易验证，当以 $q(t)f(x)$ 作为 $f(t,x)$ 时条件($H_5$)($H_6$)满足，于是用 2.6.4 得出 BVP(25)有正解. 同理，条件($H_8$)亦推出(25)有正解.

**注** 若 $\alpha_1=0,\beta_1=\infty$，则条件($H_7$)必满足；若 $\beta_2=0,\alpha_2=\infty$，则条件($H_8$)必满足.

# 第三章　单调算子

单调实函数有两种很不相同的推广：其一是有序 Banach 空间中的增(或减)映射，这已在上章中考虑；其二是本章中讨论的单调算子，这类算子在微分与积分方程问题、凸分析、凸最优化及变分不等式等领域有极重要的应用.单调算子的最重要的性质之一是它在某种“强制”条件下的满射性，正是这一性质构成单调算子在方程可解性等问题上的应用的基础.

## §1　单调算子概念

鉴于只有在多值算子的框架内，单调算子理论才能得到完全的展开，首先概述有关多值算子的基本用语，这些用语亦为后面几章所需.

给定非空集 $T,S$，称任何映射 $F:T\to 2^S$ 为从 $T$ 到 $S$ 的多值(或集值)算子，称

$$\mathrm{Gr}F \triangleq \{(x,y)\in T\times S: y\in Fx\} \tag{1}$$

为 $F$ 的图形.显然 $F$ 与 $\mathrm{Gr}F$ 相互唯一确定，因此常将二者视为等同.任给 $A\subset T$，令 $FA=\bigcup\limits_{x\in A}Fx$.称 $R(F)\triangleq FT$ 为 $F$ 的值域，称 $D(F)\triangleq\{x\in T: Fx\neq\varnothing\}$ 为 $F$ 的定义域.当写出 $F:D\subset T\to 2^S$ 时意味着 $D=D(F)$ 且 $D\neq\varnothing$. $F$ 的逆 $F^{-1}:S\to 2^T$ 定义为 $F^{-1}y=\{x: y\in Fx\}$ $(\forall\ y\in S)$.显然 $D(F^{-1})=R(F)$，$R(F^{-1})=D(F)$.所宜注意者，多值算子的逆恒可定义，这是多值算子的主要优点之一.设 $A\subset T$，“$F$ 在 $A$ 上的限制”$F|A$ 由 $(F|A)(x)=Fx\,(x\in A)$ 定义.若 $\mathrm{Gr}F\subset\mathrm{Gr}F_1$，则说 $F_1$ 是 $F$ 的扩张，写作 $F\subset F_1$.

关于 $F:T\to 2^S$ 的进一步的概念涉及 $T$ 与 $S$ 的一定(拓扑或度

量)结构. 主要定义汇集于下:

**3.1.1 定义** 设 $F:T\to 2^S, x\in T$.

1° 设 $T,S$ 为度量空间. 若对任给有界集 $A\subset T$, $FA$ 有界, 则说 $F$ 有界; 若 $\forall\ x\in T$, 存在 $x$ 的邻域 $V\subset T$: $FV$ 有界, 则说 $F$ 局部有界.

2° 设 $T,S$ 是拓扑空间. 若对任何网 $\{(x_\alpha,y_\alpha)\}\subset \mathrm{Gr}F$: $(x_\alpha,y_\alpha)\to(x,y)\Rightarrow y\in Fx$, 则说 $F$ 在 $x$ 闭. 若 $\mathrm{Gr}F$ 在 $T\times S$ 中闭, 则称 $F$ 为闭算子. 若 $T\subset X, S=Y^*$, 当 $X,Y^*$ 中分别用强拓扑与弱$^*$拓扑时 $F$ 闭, 则说 $F$ 为 sw$^*$-闭; sw-闭或 ws-闭仿此.

3° 设 $T$ 是拓扑空间, $S=Y$. 若 $\forall\ \varepsilon>0$, 存在 $x$ 的邻域 $U$: $FU\subset Fx+B_\varepsilon(0)$, 则说 $F$ 在 $x$ 上半连续(缩写为 usc). 若 $T\subset X, S=Y^*$, 对 $Fx$ 的任给弱$^*$邻域 $V$, $\exists\ \delta>0$: $F(T\cap B_\delta(x))\subset V$, 则说 $F$ 在 $x$ 为 sw$^*$-上半连续(缩写为 sw$^*$-usc); sw-上半连续(sw-usc)仿此. 若 $\forall\ x\in D(F)$: $F$ 在 $x$ 为 usc[sw$^*$-usc], 则说 $F$ 是 usc[sw$^*$-usc]的.

若 $F$ 是单值的(即 $\forall\ x\in D(F)$, $Fx$ 恰含一元), 则依 3.1.1 的 usc 亦即连续, 而 sw$^*$-usc 重合于下面 3.1.5 所定义的"次连续".

设 $F:D\subset X\to 2^{X^*}$, $G:D\subset X\to X^*$.

**3.1.2 定义** 若 $\forall\ x,y\in D$: $\langle Fx-Fy,x-y\rangle\geqslant 0$ [$x\neq y\Rightarrow\langle Fx-Fy,x-y\rangle>0$], 则称 $F$ 为单调[严格单调]算子. 若 $\exists\ \lambda>0$, $\forall\ x,y\in D$: $\langle Fx-Fy,x-y\rangle\geqslant\lambda|x-y|^2$, 则说 $F$ 强单调, 或说 $F$ 为 $\lambda$-强单调.

**3.1.3 例** 1° 设 $\varphi: D\subset\mathbf{R}\to\mathbf{R}$, 则 $\varphi$[严格]增加 $\Longleftrightarrow\varphi$ 依 3.1.2 [严格]单调. 若 $\varphi$ 增加, 则 $\tilde{\varphi}: x\mapsto[\varphi(x-0),\varphi(x+0)]$ 是单调算子且 $\varphi\subset\tilde{\varphi}$. 若 $\varphi\in C^1[a,b]$, $\varphi'(x)>0$, 则由中值定理推出 $\varphi$ 依 3.1.2 是强单调的.

以上事实表明单调算子是单调增函数的推广. 实际上, 不难证明, $G:X\to X^*$ 单调 $\Longleftrightarrow\forall\ x,y\in X$: $t\mapsto\langle G(x+ty),y\rangle$ $(t\in J)$ 为增函数.

2° 线性算子 $A:X\to X^*$ 单调、严格单调与 $\lambda$-强. 单调分别意味着 $\langle Ax,x\rangle\geqslant 0$, $\langle Ax,x\rangle>0$ 与 $\langle Ax,x\rangle\geqslant\lambda|x|^2(\forall\ x\in X\backslash\{0\})$；在这三种情况下，分别称 $A$ 为正算子，严格正算子与强正算子(注意与 §2.3 意义下的正线性算子相区别!). 对于 $A=A^*\in L(\mathbf{R}^n)$, $A$ 正 $\Leftrightarrow A$ 半正定；$A$ 严格正 $\Leftrightarrow A$ 强正 $\Leftrightarrow A$ 正定.

3° 设 $f\in C^1(\Omega,X^*)$, $\Omega\subset X$ 是凸开集，则

$$\langle f(x)-f(y),x-y\rangle=\int_0^1\langle f'(y+t(x-y))(x-y),x-y\rangle\mathrm{d}t.$$

据此易推出：$f$ 单调 $\Leftrightarrow\forall\ x\in\Omega:f'(x)$ 为正算子；$f'(x)(\forall\ x\in\Omega)$ 严格正 $\Rightarrow f$ 严格单调.

在一定意义上，单调算子可看作是正线性算子的非线性推广.

下面考虑单调算子的某些基本性质. 首先，由定义直接看出：强单调 $\Rightarrow$ 严格单调 $\Rightarrow$ 单调；$F$ 单调 $\Leftrightarrow F^{-1}$ 单调；$F$ 严格单调 $\Rightarrow F^{-1}$ 单值；$G$ 严格单调 $\Leftrightarrow G^{-1}$ 严格单调 $\Rightarrow G$ 与 $G^{-1}$ 皆为单射. 其次，对单调算子的有界性与连续性有一些较强结论.

**3.1.4 命题** 若 $F$ 单调，则 $F|D^\circ$ 局部有界.

**证** 任给 $x_0\in D^\circ$，今证 $\exists\ r>0:FB_r(x_0)$ 有界. 不妨设 $x_0=0$，且 $S\triangleq\overline{B}_\rho(0)\subset D$, $\rho>0$. 令

$$S_n=\{y\in S:\langle Fx,y-x\rangle\leqslant n(\forall\ x\in S)\},$$

则 $S_n$ 闭且易验知 $S=\bigcup_1^\infty S_n$. 由 Baire 定理，必有某个 $S_n$ 含一个球 $B_{3r}(z)$, $r>0$. 设 $-z\in S_m$，则对任给 $x\in B_r(0)$, $u\in Fx$ 有

$$|u|=\sup_{|y|=r}\frac{1}{r}u(y)=\sup_{|y|=r}\frac{1}{r}[u(-z-x)+u(y+z+2x-x)]$$
$$\leqslant\frac{m+n}{r}.$$

□

由 3.1.4 推出，若 $A:X\to X^*$ 是单调线性算子，则 $A$ 连续.

**3.1.5 定义** 设 $x\in D$. 若 $\forall\ z\in X$，当 $t\downarrow 0$ 时 $G(x+tz)\rightharpoonup Gx$，则说 $G$ 在 $x$ 半连续(hemicontinuous)；若 $x_n\to x\Rightarrow Gx_n\rightharpoonup Gx$，则说 $G$ 在 $x$ 次连续(demicontinuous). $G$ 在 $D$ 上半[次]连续 $\Leftrightarrow G$ 在每点 $x\in D$ 半[次]连续.

**3.1.6 命题** 设 $X$ 自反，$G$ 单调、半连续，则 $G|D^\circ$ 次连续，因此当 $X=\mathbf{R}^n$ 时 $G|D^\circ$ 连续.

**证** 设 $x_n \to x \in D^\circ$，则 $\{Gx_n\}$ 有界(3.1.4)，因此其任何子列含弱 * 收敛子列. 不妨设 $Gx_n \xrightarrow{*} u$，则

$$\langle Gx-u,z\rangle=\lim_{t\downarrow 0}\lim_{n}\frac{1}{t}\langle G(x+tz)-Gx_n,tz\rangle$$

$$\geqslant\lim_{t\downarrow 0}\lim_{n}\frac{1}{t}\langle G(x+tz)-Gx_n,x_n-x\rangle=0(\forall z\in X),$$

可见 $Gx=u$，从而 $Gx_n \xrightarrow{*} Gx$，即 $G$ 在 $x$ 次连续. □

**3.1.7 定义** (i)若 $F$ 单调，且当 $(y,v)\in X\times X^*$ 满足 $\langle Fx-v,x-y\rangle\geqslant 0(\forall x\in D)$ 时 $v\in Fy$，则说 $F$ 极大单调. (ii)若从 $x_n\rightharpoonup x$ 与 $\overline{\lim_{n}}\langle Gx_n,x_n-x\rangle\leqslant 0$ 恒推出 $\langle Gx,x-y\rangle\leqslant\underline{\lim_{n}}\langle Gx_n,x_n-y\rangle(\forall y\in D)$，则称 $G$ 为伪单调算子.

**注** 1° 不难验证，单调算子 $F$ 为极大单调$\Longleftrightarrow F$ 没有自身以外的单调扩张.

2° 伪单调性的定义初看起来有些古怪，但事实证明它不失为一有用的概念. 若 $G:\mathbf{R}^n\to\mathbf{R}^n$，则显然 $G$ 连续$\Rightarrow G$ 伪单调；$G$ 局部有界且伪单调$\Rightarrow G$ 连续. 这表明伪单调未必单调.

极大单调算子有一些较理想的性质.

**3.1.8 命题** 对极大单调算子 $F$ 有以下结论：(i)每个 $Fx$ 弱*闭凸；(ii)$F$ 是 sw*-闭与 ws-闭算子；(iii)若 $X$ 自反，则 $F|D^\circ$ 为 sw-usc，当 $F$ 单值时 $F|D^\circ$ 次连续；(iv)若 $X$ 自反，则 $F^{-1}$ 亦极大单调.

**证** (i)由等式 $Fx=\bigcap_{(y,v)\in GrF}\{u:\langle u-v,x-y\rangle\geqslant 0\}$ 推出.

(ii)与(iv)的证明是直接的.

(iii)若 $F$ 在某点 $x\in D^\circ$ 非 sw-usc，则有 $Fx$ 的弱邻域 $V$，$x_n\to x$，$\forall n\geqslant 1$，$\exists u_n\in(Fx_n)\setminus V$. $u_n$ 必有界，不妨设 $u_n\rightharpoonup u$，于是由 $F$ 为 sw-闭有 $u\in Fx$；而由 $u_n\rightharpoonup u$ 得 $u_n\in V$($n$ 充分大)，得出矛盾. □

**3.1.9 定理** 设 $G:X\to X^*$ 单调. (i)若 $G$ 半连续，则 $G$ 极大单

调.(ii)若 $X$ 自反,则 $G$ 极大单调$\Longleftrightarrow G$ 次连续$\Rightarrow G$ 伪单调.

**证** (i)设$\langle Gx-v,x-y\rangle\geqslant 0(\forall\ x\in X)$,取 $x=y+tz$ 并令 $t\downarrow 0$ 得$\langle Gy-v,z\rangle\geqslant 0(\forall\ z\in X)$,因此 $Gy=v$.

(ii)只需证 $G$ 次连续$\Rightarrow G$ 伪单调. 设 $x_n\to x,\overline{\lim_n}\langle Gx_n,x_n-x\rangle\leqslant 0$. 这结合

$$\underline{\lim_n}\langle Gx_n,x_n-x\rangle\geqslant\lim_n\langle Gx,x_n-x\rangle=0$$

得出$\langle Gx_n,x_n-x\rangle\to 0$. $\forall\ y\in X$,令 $y_k=(1-k^{-1})x+k^{-1}y$,则 $y_k\to x$,由 $G$ 次连续有 $Gy_k\to Gx;k(x-y_k)=x-y$. 于是

$$\begin{aligned}\langle Gx,x-y\rangle&=\lim_k\lim_n k\langle Gy_k,x_n-y_k\rangle\\&\leqslant\underline{\lim_k}\ \underline{\lim_n}k\langle Gx_n,x_n-y_k\rangle\\&=\underline{\lim_k}\ \underline{\lim_n}k\langle Gx_n,x-y_k\rangle\\&=\underline{\lim_n}\langle Gx_n,x-y\rangle=\underline{\lim_n}\langle Gx_n,x_n-y\rangle.\end{aligned}$$ □

**3.1.10 例** 1° 设 $\varphi:\mathrm{R}\to\mathrm{R}$ 为增函数,$\tilde{\varphi}$依 3.1.3 之 1°,则易直接验知 $\tilde{\varphi}:\mathrm{R}\to 2^{\mathrm{R}}$ 极大单调.

2° 由 3.1.9,任何连续单调算子 $G:X\to X^*$ 是极大单调的. 特别,任何线性单调算子 $A:X\to X^*$ 极大单调(当 $X$ 自反时 $A^{-1}$亦极大单调).

## §2 对偶映射

对偶映射是单调算子的重要例子,它对于单调算子理论的展开是不可缺少的. 半内积与对偶映射密切相关,本节一并予以讨论.

**3.2.1 定义** 任给 $x,y\in X$,令

$$(x,y)_+=\lim_{t\downarrow 0}\frac{1}{t}|y|(|y+tx|-|y|),\tag{1}$$

$(x,y)_-=-(-x,y)_+$,称$(x,y)_\pm$为半内积. 称由

$$\mathscr{D}x=\{u\in X^*;u(x)=|x|^2=|u|^2\}\tag{2}$$

定义的映射 $\mathscr{D}:X\to 2^{X^*}$ 为 $X$ 的对偶映射.

若令 $\varphi(x)=|x|$,则(1)表明

$$(x,y)_{\pm}=|y|D_{\pm}\varphi(y,x),\ x,y\in X. \tag{3}$$

因 $\varphi$ 是凸函数,故由(3)知 $(x,y)_{\pm}$ 恒有定义且 $(x,y)_{-}\leqslant(x,y)_{+}$ (参考 4.1.3). 其次,不难由 Hahn-Banach 定理推出 $\mathscr{D}x\neq\varnothing(\forall\ x\in X)$. 若 $X$ 是 Hilbert 空间,则 $(x,y)_{\pm}=(x,y)$,$\mathscr{D}=I$. $(\cdot,\cdot)_{\pm}$ 与 $\mathscr{D}$ 可分别看作是 Hilbert 空间中的内积与单位映射在 Banach 空间中的替代;空间 $X$ 愈接近 Hilbert 空间,替代的效果就愈好.

关于 $(\cdot,\cdot)_{\pm}$ 与 $\mathscr{D}$ 的一些基本性质汇集于下.

**3.2.2 命题** 对任给 $x,y,z\in X$ 以下结论成立:(i) $|(x,y)_{\pm}|\leqslant|x||y|$,$(x,x)_{\pm}=|x|^2$. (ii) $(x,y)_{-}+(z,y)_{\pm}\leqslant(x+z,y)_{\pm}\leqslant(x,y)_{\pm}+(z,y)_{+}$;$(\alpha x,\beta y)_{\pm}=\alpha\beta(x,y)_{\pm}(\alpha\beta\geqslant0)$. (iii) $(\cdot,\cdot)_{+}$ 上半连续,$(\cdot,\cdot)_{-}$ 下半连续,$(\cdot,y)_{\pm}$ 连续. (iv) $\mathscr{D}(\alpha x)=\alpha\mathscr{D}x$ $(\alpha\in\mathrm{R})$. (v) $\mathscr{D}y=\{u\in X^*:(\cdot,y)_{-}\leqslant u\leqslant(\cdot,y)_{+}\}$. (vi) $(x,y)_{+}=\max\langle\mathscr{D}y,x\rangle$,$(x,y)_{-}=\min\langle\mathscr{D}y,x\rangle$. (vii) $(x+\alpha y,y)_{\pm}=(x,y)_{\pm}+\alpha|y|^2(\alpha\in\mathrm{R})$.

**证** 只需考虑关于 $(\cdot,\cdot)_{+}$ 的结论. (i)是明显的. (ii)由(3)得出. 利用(1)可直接证明 $(\cdot,\cdot)_{+}$ 为 usc,从而 $(\cdot,y)_{+}$ 亦 usc. 若 $x_n\to x$,则利用(i) (ii)得

$$(x,y)_{+}\leqslant\varliminf_{n}(x_n,y)_{+}+\varlimsup_{n}(x-x_n,y)_{+}=\varliminf_{n}(x_n,y)_{+},$$

可见 $(\cdot,y)_{+}$ 为 lsc,从而连续. (iv)是明显的.

证(v). 若 $u\in\mathscr{D}y$,则 $tu(x)=u(y+tx)-u(y)\leqslant|y|(|y+tx|-|y|)$,这推出 $(x,y)_{-}\leqslant u(x)\leqslant(x,y)_{+}$. 反之,从 $(x,y)_{-}\leqslant u(x)\leqslant(x,y)_{+}$ 推出 $|u(x)|\leqslant|x||y|$,$u(y)=|y|^2$,由此得 $u\in\mathscr{D}y$.

证(vi). 给定 $x,y\in X$,由(v)有 $\langle\mathscr{D}y,x\rangle\leqslant(x,y)_{+}$. 定义 $u:\mathrm{R}x\to\mathrm{R}$,$\lambda x\mapsto\lambda(x,y)_{+}$,则 $u(\lambda x)\leqslant(\lambda x,y)_{+}(\forall\ x\in\mathrm{R})$. 由 Hahn-Banach 定理知有 $u\in X^*$:$(\cdot,y)_{-}\leqslant u\leqslant(\cdot,y)_{+}$,$u(\lambda x)=\lambda(x,y)_{+}$[75;1.7.1],因此 $u\in\mathscr{D}y$,$u(x)=(x,y)_{+}$. 这得出 $(x,y)_{+}=$

$\max\langle \mathscr{D}y, x\rangle$. 由(vi)直接推出(vii). □

**3.2.3 命题** $\mathscr{D}$ 是极大单调的.

**证** 单调性是明显的. 设 $(y,v)\in X\times X^*$ 满足 $\langle \mathscr{D}x-v, x-y\rangle \geqslant 0(\forall\ x\in X)$,则分别取 $x=y+tz$ 与 $x=(1\pm t)y(t>0, z\in X)$得

$$0\leqslant \langle \mathscr{D}(y+tz)-v, tz\rangle = t[\langle \mathscr{D}(y+tz), z\rangle - v(z)];$$

$$0\leqslant \langle \mathscr{D}(y\pm ty)-v, \pm ty\rangle = \pm t[(1\pm t)|y|^2 - v(y)].$$

由此两式分别推出 $|v|\leqslant |y|$ 与 $v(y)=|y|^2$,因此 $v\in \mathscr{D}y$. □

以上结果不依赖于空间 $X$ 的特殊性质,必然只能提供很一般的信息,完全不足以满足一些精细讨论之需要. 因此必要逐步增加对 $X$ 的限制,以获得关于 $(\cdot,\cdot)_+$ 与 $\mathscr{D}$ 的更强的结论.

让我们回顾一下某些 Banach 空间概念. 以下分别以 $B$ 与 $B^*$ 记 $X$ 与 $X^*$ 中的单位球. 若 $\partial B$ 不含多于一点的线段,则说 $X$ 严格凸. 若 $\forall\ x\in\partial B, \varepsilon>0, \exists\ \delta=\delta(x,\varepsilon)$,使当 $y\in\partial B$, $|x-y|\geqslant\varepsilon$时 $|x+y|\leqslant 2(1-\delta)$,则说 $X$ 局部一致凸;若上述的 $\delta(x,\varepsilon)$与 $x$ 无关,则说 $X$ 一致凸. 我们综述以下结果以备引用(有关详细讨论参看[43,198,202]).

**3.2.4 定理** (i)$X$ 严格凸$\Longleftrightarrow\forall\ u\in X^*\backslash\{0\}$:至多有一点 $x\in\partial B$ 使$u(x)=|u|$;$X^*$ 严格凸$\Rightarrow\forall\ x\in X\backslash\{0\}$:有唯一$u\in\partial B^*$ 使 $u(x)=|x|$. (ii)若 $X$ 局部一致凸,则 $X$ 严格凸,当 $x_n\rightharpoonup x\in X$ 且 $|x_n|\to|x|$时 $x_n\to x$. (iii)$X$ 一致凸$\Longleftrightarrow X^*$ 中的范数$|\cdot|$在 $\partial B^*$ 上一致可微$\Rightarrow X$ 自反;Hilbert 空间必一致凸. (iv)$X$ 自反$\Longleftrightarrow\forall\ u\in X^*, \exists\ x\in\partial B: u(x)=|u|$;当 $X$ 自反时可改赋等价范数使 $X$ 与 $X^*$ 皆成为局部一致凸空间.

粗略地说,空间凸性愈好(这意味着其中的球更类似于 Euclid 球),它就愈接近于 Hilbert 空间;因而,如下面的定理所表明的,半内积与对偶映射的性质就愈好.

**3.2.5 定理** 关于 $\mathscr{D}: X\to 2^{X^*}$ 有以下结论:(i)$X$ 严格凸$\Longleftrightarrow$ $\mathscr{D}$ 严格单调. (ii)若 $X^*$ 严格凸,则 $\mathscr{D}$ 是单值的且 $\mathscr{D}x=$

$D\left(\frac{1}{2}|x|^2\right)=|x|D|x|\,(x\neq0)$，$D$ 记 $G$-导数；$(x,y)_\pm=\langle\mathscr{D}y,x\rangle$. (iii)若 $X$ 自反，$X^*$ 严格凸，则 $\mathscr{D}:X\to X^*$ 伪单调且次连续. (iv)若 $X$ 自反，$X$ 与 $X^*$ 皆严格凸，则 $\mathscr{D}:X\to X^*$ 为双射，且 $\mathscr{D}^{-1}$是 $X^*$ 的对偶映射. (v)若 $X$ 自反，$X^*$ 局部一致凸，则 $\mathscr{D}:X\to X^*$ 连续；若 $X$ 亦局部一致凸，则 $\mathscr{D}$ 是同胚. (vi)若 $X^*$ 一致凸，则 $\mathscr{D}$ 与$(\cdot,\cdot)_+$在有界集上一致连续.

**证** 利用 3.2.4(i)，不难直接证明(i).

(ii)$\forall\, x\in X\backslash\{0\}$，$u,v\in\mathscr{D}x$，若 $u\neq v$，则由 $X^*$严格凸有 $t\in J$: $|tu+t'v|^2<|u|^2=|x|^2=\langle tu+t'v,x\rangle<|x|^2$，得出矛盾. 这表明 $\mathscr{D}$ 单值；因而由 3.2.2(vi)得$(x,y)_\pm=\langle\mathscr{D}y,x\rangle$. 令 $\varphi(x)=|x|$，则结合(3)有

$$\langle\mathscr{D}x,z\rangle=(z,x)_\pm=|x|D_\pm\varphi(x,z),$$

这表明$|x|^{-1}\mathscr{D}x=D|x|$，$\mathscr{D}x=|x|D|x|=D\left(\frac{1}{2}|x|^2\right)\,(0\neq x\in X)$.

结合(ii)，3.2.3 与 3.1.9 得出(iii).

(iv)以 $J$ 记 $X^*$ 的对偶映射为 $J$. 由(i)，$\mathscr{D}$ 与 $J$ 皆严格单调，从而皆为单射. 其次，$\mathscr{D}^{-1}$与 $J$ 皆极大单调，且易验知 $\mathscr{D}^{-1}\subset J$，故必 $\mathscr{D}^{-1}\subset J$，$\mathscr{D}$ 是双射.

(v)若 $x_n\to x$，则$|\mathscr{D}x_n|=|x_n|\to|x|=|\mathscr{D}x|$；由(iii)有 $\mathscr{D}x_n\rightharpoonup\mathscr{D}x$，于是 $\mathscr{D}x_n\to\mathscr{D}x$(3.2.4(ii))，可见 $\mathscr{D}$ 连续. 若 $X$ 亦局部一致凸，则由(iv)知 $\mathscr{D}$ 为双射且 $\mathscr{D}^{-1}$是 $X^*$ 的对偶映射，从而 $\mathscr{D}^{-1}$亦连续.

(vi)$\mathscr{D}$ 在 $\partial B$ 上必一致连续. 否则有 $\epsilon>0$，$x_n,y_n\in\partial B$: $x_n-y_n\to0$，$|\mathscr{D}x_n-\mathscr{D}y_n|\geqslant\epsilon$. 由 $X^*$ 一致凸有 $\delta>0$:

$$\begin{aligned}2(1-\delta)&\geqslant|\mathscr{D}x_n+\mathscr{D}y_n|\geqslant\langle\mathscr{D}x_n+\mathscr{D}y_n,x_n\rangle\\&=2+\langle\mathscr{D}y_n,x_n-y_n\rangle\geqslant2-|x_n-y_n|\to2\,(n\to\infty),\end{aligned}$$

得出矛盾. 考虑到 $\mathscr{D}$ 是齐次的，知 $\mathscr{D}$ 在任何有界集 $D\subset X\backslash B_r(0)$ $(r>0)$上一致连续；再注意$|\mathscr{D}x|=|x|$，知 $\mathscr{D}$ 在任何有界集上一

致连续. □

试看两个具体例子.

**3.2.6 例** 1° 设 $X=L^p(\Omega)$, $1<p<\infty$, $(\Omega,\mu)$是一测度空间. $X$ 与 $X^*$ 皆一致凸,因此 $\mathscr{D}:X\to X^*$ 是同胚. 今求出 $\mathscr{D}x$ 的表达式. 设 $0\neq x\in X$, $y\in X$,则

$$\begin{aligned}\langle |x|_p^{-1}\mathscr{D}x,y\rangle &= \frac{\mathrm{d}}{\mathrm{d}t}|x+ty|_p|_{t=0}\\ &= \frac{1}{p}|x|_p^{1-p}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega|x+ty|^p\mathrm{d}\mu|_{t=0}\\ &= |x|_p^{1-p}\int_\Omega|x|^{p-1}(\mathrm{sgn}x)y\mathrm{d}\mu,\end{aligned}$$

这得出 $\mathscr{D}x=|x|_p^{2-p}|x|^{p-2}x(0\neq x\in X)$. 注意当 $p=2$ 时 $\mathscr{D}x=x$,这正如所预期的.

2° 设 $X=C(T)$, $T$ 是一紧 $T_2$ 空间,则有公式:

$$(x,y)_+=\max_{t\in T_y}x(t)y(t),(x,y)_-=\min_{t\in T_y}x(t)y(t);\tag{4}$$

$$\mathscr{D}x=\left\{v\left|\begin{array}{c}v=v\text{ 是 }T\text{ 上的有界测度,}\\ \|v\|=|x|_0,\mathrm{supp}v\subset T_x,(\mathrm{sgn}x)v\geqslant 0\end{array}\right.\right\},\tag{5}$$

其中 $T_x=\{t\in T:|x(t)|=|x|_0\}$. 仅证(4)的第一式:$\forall t\in T_y$: $x(t)\cdot y(t)=\langle|y|_0\delta_t,x\rangle\leqslant(x,y)_+$($\delta_t$ 记集中于 $t$ 的 Dirac 测度,易验知 $|y|_0\delta_t\in\mathscr{D}y$). 其次,对 $\lambda_n\downarrow 0$,取 $t_n\in T_{y+\lambda_n x}$,可设 $t_n\to t\in T_y$, $|y|_0>0$,于是

$$\begin{aligned}|y|_0^{-1}(x,y)_+&=\lim_n\frac{1}{\lambda_n}[|y(t_n)+\lambda_n x(t_n)|-|y(t)|]\\ &\leqslant\lim_{\lambda\downarrow 0}\lim_n\frac{1}{\lambda}[|y(tn)+\lambda x(tn)|-|y(t)|]\\ &=\lim_{\lambda\downarrow 0}\frac{1}{\lambda}[|y(t)+\lambda x(t)|-|y(t)|]\\ &=|y(t)|^{-1}x(t)y(t),\end{aligned}$$

这得出 $(x,y)_+=\max\{x(t)y(t):t\in T_y\}$.

## §3 满射定理

本节设 $X$ 自反，$F: D\subset X\to 2^{X^*}$，$G: C\subset X\to X^*$.

给定 $b\in X^*$. 若存在 $x_0\in \dot{D}$，使 $b\in Fx_0$，则称 $x_0$ 为"多值算子方程"(或"包含")

$$b\in Fx \tag{1}$$

的解. 当 $F$ 单值时，(1)就是通常的算子方程 $Fx=b$. 显然(1)可解 $\Longleftrightarrow b\in R(F)$，因此当 $R(F)=X^*$ 时(1)对任何 $b\in X^*$ 可解. 可见满射性对于方程的可解性有重要意义. 简单的例子(如 $Fx=\mathrm{arctg}x$)表明，即使连续极大单调算子亦未必为满射. 不过，"强制的"极大单调算子必为满射，本节将经几步达此结论. 关键的想法是考虑(1)的"伪单调扰动"

$$b\in Fx+Gx, \tag{2}$$

(2)本身亦有重要应用价值. 关于(2)的以下结果通常被认为是单调算子理论的最重要成果之一.

**3.3.1 定理**(Browder, 1968)　设 $F$ 极大单调，$G$ 有界、次连续且伪单调，$\overline{\mathrm{co}}D\subset C$. 若存在 $(\hat{x},\hat{u})\in \mathrm{Gr}F$，使当 $x\in C$，$|x|\to\infty$ 时 $\langle Gx+\hat{u}-b, x-\hat{x}\rangle>0$，则(2)可解.

**注**　若 $C$ 有界，则认定条件"$x\in C$，$|x|\to\infty$ 时 $\langle Gx+\hat{u}-b, x-\hat{x}\rangle>0$"自动满足. 本书它处遇到类似情况时亦将用同样约定.

**证**　不妨设 $\hat{x}=0$，$b=\hat{u}=0$(否则分别以 $F(\hat{x}+\cdot)-\hat{u}$ 与 $G(\hat{x}+\cdot)+\hat{u}-b$ 代 $F$ 与 $G$). 由 $F$ 极大单调只需证 $\exists\, x_0\in C$，$\forall\, x\in D$: $\langle Fx+Gx_0, x-x_0\rangle\geqslant 0$.

1° 设 $X=\mathrm{R}^n$，$D$ 有界. 令 $K=\overline{\mathrm{co}}D$，$V_{xu}=\{y\in K: \langle u+Gy, x-y\rangle<0\}$. 若所需的 $x_0$ 不存在，则 $\{V_{xu}: (x,u)\in \mathrm{Gr}F\}$ 是紧集 $K$ 的开覆盖，于是有 $K$ 上从属于 $\{V_{xu}\}$ 的单位分解 $\{\varphi_{xu}\}$([75; 1.4.13]).

令 $f(y)=\sum\limits_{x,u}\varphi_{xu}(y)x$，$\sum\limits_{x,u}$ 表对 $(x,u)\in \mathrm{Gr}F$ 求和，则 $f\in C(K,$

$K$). 由 1.4.3 有 $z\in K: f(z)=z$. 于是

$$0 = \langle \sum_{x,u}\varphi_{xu}(z)u + Gz, f(z) - z\rangle$$

$$= \langle \sum_{x,u}\varphi_{xu}(z)(u + Gz), \sum_{y,v}\varphi_{yv}(z)(y - z)\rangle$$

$$= \frac{1}{2}\sum_{x,u,y,v}\varphi_{xu}(z)\varphi_{yv}(z)(\langle u + Gz, y - z\rangle + \langle v + Gz, x - z\rangle). \tag{3}$$

若 $\varphi_{xu}(z)\varphi_{yv}(z)\neq 0((x,u),(y,v)\in \mathrm{Gr}F)$,则 $z\in V_{xu}\cap V_{yv}$,于是

$$\langle u+Gz, y-z\rangle + \langle v+Gz, x-z\rangle$$

$$=\langle u+Gz, x-z\rangle + \langle v+Gz, y-z\rangle - \langle u-v, x-y\rangle < 0.$$

这表明(3)式右端$<0$,得出矛盾.

2°若$X=\mathrm{R}^n$,$D$无界,则$\forall\ k\geqslant 1$,$\exists\ x_k\in\overline{\mathrm{co}}D$,$\forall\ x\in D\cap B_k(0)$: $\langle Fx+Gx_k, x-x_k\rangle\geqslant 0$,特别$\langle Gx_k, -x_k\rangle\geqslant 0$. 由定理条件推出$\{x_k\}$有界,不妨设 $x_k\to x_0$,则 $x_0$ 合于所求.

3° 设 $\dim X=\infty$. 以 $\sigma$ 记 $X'$的有限维子空间之全体. $\forall\ S\in\sigma$,由上段结论知$\exists$ $x_S\in S\cap\overline{\mathrm{co}}D$,$\forall\ x\in D\cap S$: $\langle Fx+Gx_S, x-x_S\rangle\geqslant 0$,特别$\langle Gx_S, -x_S\rangle\geqslant 0$,从而$\{x_S\}$与$\{Gx_S\}$有界. 令 $M_S=\{(x_T, Gx_T): S\subset T\in\sigma\}$,显然$\{M_S: S\in\sigma\}$有限相交,于是必有$(x_0,u_0)\in\bigcap_{S\in\sigma}\overline{M}_S$,$\overline{M}_S$ 记 $M_S$ 的弱闭包. 取定$(x,u)\in\mathrm{Gr}F$,今证$\langle u+Gx_0, x-x_0\rangle\geqslant 0$. 必有$(y,v)\in\mathrm{Gr}F$: $\langle v+u_0, y-x_0\rangle\leqslant 0$(否则$\forall\ y\in D$: $\langle Fy+u_0, y-x_0\rangle>0$,这得出 $x_0\in D$,令 $y=x_0$ 得出矛盾). 取 $S\in\sigma$,使 $x,y\in S$. 因$(x_0,u_0)\in\overline{M}_S$,必有$(x_n,Gx_n)\in M_S$: $x_n\rightharpoonup x_0$, $Gx_n\rightharpoonup u_0$(参考[202; Ⅰ, p. 911]). 由 $M_S$ 的定义有$\langle Fy+Gx_n, y-x_n\rangle\geqslant 0$,从而$\langle Gx_n, x_n-y\rangle\leqslant\langle v, y-x_n\rangle$,于是

$$\overline{\lim_n}\langle Gx_n, x_n-x_0\rangle\leqslant\lim_n(\langle v, y-x_n\rangle+\langle Gx_n, y-x_0\rangle)$$

$$=\langle v+u_0, y-x_0\rangle\leqslant 0.$$

由$\langle Fx+Gx_n, x-x_n\rangle\geqslant 0$ 得$\langle Gx_n, x_n-x\rangle\leqslant\langle u, x-x_n\rangle$. 因 $x_0\in\overline{\mathrm{co}}D\subset C$,故由 $G$ 伪单调得出所要的不等式:

$$\langle u+Gx_0,x-x_0\rangle=\lim_n\langle u,x-x_n\rangle-\langle Gx_0,x_0-x\rangle$$
$$\geqslant\varliminf_n\langle Gx_n,x_n-x\rangle-\langle Gx_0,x_0-x\rangle\geqslant 0. \quad \square$$

3.3.1 有一系列重要推论. 为叙述方便,引进

**3.3.2 定义** 若 $\lim_{x\in C,|x|\to\infty}\langle Gx,x\rangle/|x|=\infty$,则说 $G$ 强制;若 $\lim_{x\in C,|x|\to\infty}|Gx|=\infty$,则说 $G$ 弱强制.

显然强制⇒弱强制;若 $\varlimsup_{x\in C,|x|\to\infty}|Gx|/|x|<\infty$,则 $G$ 强制$\Longleftrightarrow$ $\forall\ x\in X$:$G(x+\cdot)$强制;若 $G(\hat{x}+\cdot)$强制,则

$$\langle Gx+\hat{u}-b,x-\hat{x}\rangle\geqslant|x-\hat{x}|\left(\frac{\langle Gx,x-\hat{x}\rangle}{|x-\hat{x}|}-|\hat{u}-b|\right)\to\infty$$

$(x\in C,|x|\to\infty)$. 以上事实结合 3.3.1 推出:

**3.3.3 推论** 设 $F$ 极大单调,$G$ 有界、次连续、伪单调且 $\exists\ \hat{x}\in D$:$G(\hat{x}+\cdot)$强制,$\overline{\text{co}}D\subset C$,则 $R(F+G)=X^*$.

在 3.3.3 中取 $F=0$ 得到:

**3.3.4 推论** 若 $G:X\to X^*$ 有界、次连续、伪单调且 $\exists\ \hat{x}\in X$:$G(\hat{x}+\cdot)$强制,则 $GX=X^*$;若 $G\in C(\mathrm{R}^n,\mathrm{R}^n)$,$\exists\ \hat{x}\in\mathrm{R}^n$:$G(\hat{x}+\cdot)$强制,则 $G\mathrm{R}^n=\mathrm{R}^n$.

以下 $\mathscr{D}$ 记 $X$ 的对偶映射.

**3.3.5 定理** 设 $X$ 与 $X^*$ 严格凸. (i)$F$ 极大单调$\Longleftrightarrow\forall\ \lambda>0$ (或$\exists\ \lambda>0$):$R(F+\lambda\mathscr{D})=X^*$. (ii)若 $F$ 极大单调,$\lambda>0$,则"预解式"$R_\lambda\triangleq(F+\lambda\mathscr{D})^{-1}:X^*\to D$ 单值且极大单调.

**证** (i)$\forall\ \lambda>0$,$G\triangleq\lambda\mathscr{D}$ 显然满足 3.3.3 之条件(参考 3.2.5),因此当 $F$ 极大单调时 $R(F+\lambda\mathscr{D})=X^*$. 反之,若$\exists\ \lambda>0$: $R(F+\lambda\mathscr{D})=X^*$,$\langle Fx-v,x-y\rangle\geqslant 0(\forall\ x\in D)$,则取 $x\in R_\lambda(v+\lambda\mathscr{D}y)$代入得$\langle\mathscr{D}y-\mathscr{D}x,x-y\rangle\geqslant 0$;由 $\mathscr{D}$ 严格单调(3.2.5(i))得 $x=y$,$v\in Fy$,可见 $F$ 极大单调.

(ii)由 $\mathscr{D}$ 严格单调推出 $F+\lambda\mathscr{D}$ 严格单调,从而 $R_\lambda$ 单值. 由已证之(i)有 $D(R_\lambda)=X^*$. 余下只需证 $R_\lambda$ 次连续(参考 3.1.9). 设 $x_n=R_\lambda u_n$,$x=R_\lambda u$,$u_n\to u$. 因$\{x_n\}$有界(3.1.4),不妨设 $x_n\rightharpoonup y$. 设 $u_n=$

$v_n+\lambda\mathscr{D}x_n, u=v+\lambda\mathscr{D}x, v_n\in Fx_n, v\in Fx$，则

$0\leqslant\langle u_n-u, x_n-x\rangle=\langle v_n-v, x_n-x\rangle+\lambda\langle\mathscr{D}x_n-\mathscr{D}x, x_n-x\rangle\to 0$,

这推出$\langle\mathscr{D}x_n-\mathscr{D}x, x_n-x\rangle\to 0$. 进而由

$$\langle\mathscr{D}x_n-\mathscr{D}x, x_n-x\rangle\geqslant(|x_n|-|x|)^2+|x||x_n|-\langle\mathscr{D}x, x_n\rangle$$

推出$|x_n|\to|x|$，$\langle\mathscr{D}x, x_n\rangle\to|x|^2=\langle\mathscr{D}x, y\rangle$，因此$|x|\leqslant|y|\leqslant\varliminf_n|x_n|=|x|$，即$|x|=|y|$. 注意到$\mathscr{D}^{-1}$是$X^*$的对偶映射(3.2.5(iv))，$\langle\mathscr{D}x, y\rangle=|\mathscr{D}x|^2=|y|^2$，得$y=\mathscr{D}^{-1}\mathscr{D}x=x$，因此$x_n\rightharpoonup x$. 这证得$R_\lambda u_n\rightharpoonup R_\lambda u$，即$R_\lambda$次连续. □

**注** 若$X$是 Hilbert 空间，$A\in L(X)$是正的，则由 3.3.5(ii)得出$\lambda<0\Rightarrow A-\lambda I\in GL(X)$. 由此得熟知结论：正自伴算子只有非负谱值.

利用 3.3.5 可得极大单调算子满射性的完全刻划.

**3.3.6 定理**(Browder, 1968) 设$F$极大单调. 则$R(F)=X^*\Longleftrightarrow F^{-1}$在$X^*$上局部有界；$D=X\Longleftrightarrow F$在$X$上局部有界.

**证** 显然只需证前一半，且只需假定$F^{-1}$在$X^*$上局部有界证$FD$在$X^*$中既开又闭. 若$u_n\in Fx_n, u_n\to u$，则$\{x_n\}$有界，不妨设$x_n\rightharpoonup x$，于是$u\in Fx$(3.1.8). 可见$FD$闭. 其次证$FD$开. 即$\forall\, x_0\in D$，$u_0\in Fx_0$，$\exists\ r>0$：$B_r(u_0)\subset FD$. 不妨设$x_0=0$. 取$r>0$，使$F^{-1}B_{2r}(u_0)$有界. 不妨设$X$与$X^*$严格凸(3.2.4). $\forall\ u\in B_r(u_0)$，$\lambda\in(0,1)$，令$x_\lambda=R_\lambda u$，即$u=u_\lambda+\lambda\mathscr{D}x_\lambda, u_\lambda\in Fx_\lambda$(3.3.5)，则

$$\lambda|x_\lambda|^2=\langle\lambda\mathscr{D}x_\lambda, x_\lambda\rangle=\langle u-u_\lambda, x_\lambda\rangle\leqslant\langle u-u_0, x_\lambda\rangle\leqslant r|x_\lambda|;$$

$$|u_\lambda-u_0|<|u_\lambda-u|+r=\lambda|x_\lambda|+r\leqslant 2r,$$

可见$u_\lambda\in B_{2r}(u_0)$，从而$x_\lambda\in F^{-1}B_{2r}(u_0)$，于是$x_\lambda$有界. 这推出$u_\lambda\to u(\lambda\downarrow 0)$；由$FD$闭有$u\in FD$，故$B_r(u_0)\subset FD$得证. □

由 3.3.6 可得出以下更具体的满射性结果.

**3.3.7 推论**(Browder, 1963) 若$F$极大单调，且弱强制，即$\lim\limits_{x\in D, |x|\to\infty}d(0, Fx)=\infty$，则$R(F)=X^*$.

**证** 所给条件保证$F^{-1}$在$X^*$上局部有界. □

结合 3.3.7 与 3.1.9 得出：

**3.3.8 推论** 若 $G: X\to X^*$ 单调、半连续且弱强制，则 $GX=X^*$；若再假定 $G$ 严格单调[强单调]，则 $G^{-1}: X^*\to X$ 单值、严格单调且次连续[为 Lip].

**注** 1°若 $G:\mathbb{R}^n\to\mathbb{R}^n$ 为单调满射，则必为弱强制.

2° 若 $X$ 非自反，则 $\exists\ u\in X^*$，$\forall\ x\in\partial B_1(0): u(x)<|u|$（3.2.4(iv)）；显然 $u\overline{\in}R(\mathscr{D})$. 可见 3.3.7 在非自反空间中不真.

# §4 Yosida 逼近与和定理

本节设 $X$ 与 $X^*$ 自反、严格凸，$F: D\subset X\to 2^{X^*}$ 极大单调，$\mathscr{D}$ 记 $X$ 的对偶映射. 本节的课题是用一个具较好性质的单值算子 $F_\lambda$ 来"逼近"$F$，并通过这种逼近来获得有关 $F$ 的较深入的结果.

**3.4.1 引理** 任给 $\lambda>0$，$x\in X$，存在唯一 $(y,v)\in \mathrm{Gr}F$，使 $\mathscr{D}(y-x)+\lambda v=0$.

**证** 以 $\lambda F(x+\cdot)$ 代 $F$ 应用 3.3.5(ii)，知有唯一 $z\in X$ 使 $0\in\mathscr{D}z+\lambda F(x+z)$. 令 $y=x+z$，$v=-\lambda^{-1}\mathscr{D}z$，则 $(y,v)$ 如引理所求且是唯一的. □

分别以 $R_\lambda x$ 与 $F_\lambda x$ 记 3.4.1 中的 $y,v$，则得到映射 $R_\lambda: X\to D$ 与 $F_\lambda: X\to X^*$，它们由关系

$$\mathscr{D}\circ(I-R_\lambda)=\lambda F_\lambda, F_\lambda\subset F\circ R_\lambda \tag{1}$$

唯一确定且彼此相互确定. 称 $F_\lambda$ 为 $F$ 的"Yosida 逼近". 如下面即将指明的，当 $\lambda\downarrow 0$ 时 $F_\lambda|D$ 逼近于 $F$ 的一个选择 $F_0: D\to X^*$，$F_0$ 决定于条件：

$$F_0x\in Fx, |F_0x|=d(0,Fx). \tag{2}$$

因 $Fx$ 弱闭凸而 $X^*$ 严格凸，(2)唯一确定 $F_0x$. 若 $X$ 是 Hilbert 空间，则 $R_\lambda=(I+\lambda F)^{-1}$，这与 3.3.5 中的 $R_\lambda$ 相当.

**3.4.2 定理** $R_\lambda$ 与 $F_\lambda(\lambda>0)$ 有以下性质：(i)$\forall\ x\in\overline{D}: R_\lambda x\to x$ $(\lambda\downarrow 0)$. (ii)$\forall\ x\in D: \lambda^{-1}|x-R_\lambda x|=|F_\lambda x|\leqslant|F_0x|$. (iii)$\forall\ x\in D$:

$F_\lambda x \rightharpoonup F_0 x(\lambda\downarrow 0)$（若$X^*$局部一致凸则$F_\lambda x\to F_0 x$）；$\forall\ x\in X\backslash\overline{D}$：$|F_\lambda x|\to\infty(\lambda\downarrow 0)$. (iv)$F_\lambda$ 有界且极大单调.

**证** 设 $x\in X, y\in D$，令 $x_\lambda = x - R_\lambda x$. 由

$$|x_\lambda|^2 = \langle\mathscr{D}x_\lambda, x_\lambda\rangle = \langle\mathscr{D}x_\lambda, x-y\rangle + \langle\lambda F_\lambda x, y - R_\lambda x\rangle$$
$$\leqslant\langle\mathscr{D}x_\lambda, x-y\rangle + \inf\langle\lambda F y, y - R_\lambda x\rangle$$

得
$$|x_\lambda|^2 \leqslant \langle\mathscr{D}x_\lambda, x-y\rangle + \lambda|F_0 y||y - R_\lambda x|; \tag{3}$$
$$|x_\lambda|^2 \leqslant |x_\lambda||x-y| + \lambda|F_0 y|(|y-x| + |x_\lambda|). \tag{4}$$

(4)表明当 $\lambda\downarrow 0$ 时 $x_\lambda$ 有界. 任给 $\lambda_n\downarrow 0$，不妨设 $\mathscr{D}x_{\lambda_n}\rightharpoonup u$，于是由(3)有

$$\overline{\lim_n}|x_{\lambda_n}|^2 \leqslant \langle u, x-y\rangle, \tag{5}$$

因(5)亦适用于$y\in\overline{\mathrm{co}}D$，故取$x=y\in\overline{\mathrm{co}}D$得$x_{\lambda_n}\to 0$，因此$R_\lambda x\to x(\lambda\downarrow 0)$. 由此也得出$\overline{\mathrm{co}}D=\overline{D}$. 若 $x\in D$，则在(3)中取 $y=x$ 得$|x_\lambda|\leqslant\lambda|F_0 x|$. 这就证得(i)(ii).

证(iii). 设 $x\in D$. 任给 $\lambda_n\downarrow 0$，由(ii)有$|F_{\lambda_n}x|\leqslant|F_0 x|$，故不妨设 $F_{\lambda_n}x\rightharpoonup u$. 任给$(y,v)\in\mathrm{Gr}F$，有

$$\langle u-v, x-y\rangle = \lim_n\langle F_{\lambda_n}x - v, R_{\lambda_n}x - y\rangle \geqslant 0,$$

于是由 $F$ 极大单调得 $u\in Fx$；这与$|u|\leqslant|F_0 x|$及(2)一起得出 $u=F_0 x$，因此 $F_{\lambda_n}x\rightharpoonup F_0 x$，从而 $F_\lambda x\rightharpoonup F_0 x(\lambda\downarrow 0)$. 由

$$|F_0 x| \leqslant \underline{\lim_{\lambda\downarrow 0}}|F_\lambda x| \leqslant \overline{\lim_{\lambda\downarrow 0}}|F_\lambda x| \leqslant |F_0 x|$$

得$|F_\lambda x|\to|F_0 x|(\lambda\downarrow 0)$，因此当 $X^*$ 局部一致凸时 $F_\lambda x\to F_0 x(\lambda\downarrow 0)$. 若 $x\in X\backslash\overline{D}$，则$|F_\lambda x| = \lambda^{-1}|x - R_\lambda x|\geqslant\lambda^{-1}\mathrm{d}(x,D)\to\infty(\lambda\downarrow 0)$.

证(iv). 由(4)看出 $I-R_\lambda$ 有界，因此 $F_\lambda$ 有界. 可验知 $F_\lambda=(F^{-1}+\lambda\mathscr{D}^{-1})^{-1}$，故 $F_\lambda$ 极大单调(3.2.5, 3.3.5). □

3.4.2(i)的证明实际上得到了：

**3.4.3 推论** $\overline{D}$(从而$\overline{R(F)}$)为凸集.

利用以上结论，可对 $F$ 的极大单调性作出最简单的初步检定.

下面设 $F:D\subset X\to 2^{X^*}$ 与 $G:C\subset X\to 2^{X^*}$ 皆极大单调，$C\cap D\neq\varnothing$. 凡断言 $F+G$ 极大单调或为满射的结果称为"和定理"，这类定理在单调算子的应用中有基本的重要性. 下面的讨论基于如下想法：$F_\lambda+G$ 提供了 $F+G$ 的一个"近似".

**3.4.4 引理** $F_\lambda+G(\lambda>0)$ 极大单调.

**证** 取充分大的 $\mu>0$，令 $H=F_\lambda+\mu\mathscr{D}$，只要证 $R(G+H)=X^*$(3.3.5). 由 3.4.2 与 3.3.5，$H$ 有界且极大单调，从而次连续、伪单调(3.1.9). 由(4)可推出 $|F_\lambda x|\leqslant\alpha|x|+\beta,\alpha,\beta>0$ 与 $x$ 无关. 于是

$$\langle Hx,x\rangle\geqslant(\mu-\alpha)|x|^2-\beta|x|,|Hx|\leqslant(\mu+\alpha)|x|+\beta,$$

因此可用 3.3.3 得出 $R(G+H)=X^*$. □

**3.4.5 定理** 当以下条件之一满足时 $F+G$ 极大单调：(i) $C^\circ\cap D\neq\varnothing$ 或 $C\cap D^\circ\neq\varnothing$；(ii) $X$ 是 Hilbert 空间，$\forall\lambda>0,x\in C$：$(F_\lambda x,Gx)\geqslant 0$.

**证** 仅就 $X$ 为 Hilbert 空间的情况给出证明，完全的证明参看[202; Ⅱ/B]. 由 3.3.5，只要证 $R(F+G+I)=X$. 取定 $u\in X$ 令 $y_\lambda=(F_\lambda+G+I)^{-1}u$；取 $w_\lambda\in Gy_\lambda$，使 $u=F_\lambda y_\lambda+w_\lambda+y_\lambda$.

1° 证 $F_\lambda y_\lambda$ 有界，为此只要证 $y_\lambda,w_\lambda$ 有界. 不妨设 $0\in C\cap D$，$0\in F(0)\cap G(0)$，于是 $F_\lambda(0)=0$；

$$\begin{aligned}|y_\lambda|^2&=(u-F_\lambda y_\lambda-w_\lambda,y_\lambda)\\&\leqslant(u,y_\lambda)\leqslant|u||y_\lambda|;\\(w_\lambda,y_\lambda)&=(u-F_\lambda y_\lambda-y_\lambda,y_\lambda)\leqslant(u,y_\lambda),\end{aligned}$$

由此得出

$$|y_\lambda|\leqslant|u|,(w_\lambda,y_\lambda)\leqslant|u|^2(\lambda>0).\tag{6}$$

若条件(i)满足，不妨设 $0\in C^\circ$. 取 $\delta,\rho>0$，使 $B_\delta(0)\subset C$，$GB_\delta(0)\subset B_\rho(0)$(3.1.4). $\forall x\in B_\delta(0),w\in Gx$，有

$$\begin{aligned}(w_\lambda,x)&=(w,x)+(w_\lambda-w,y_\lambda)-(w_\lambda-w,y_\lambda-x)\\&\leqslant(w,x)+(w_\lambda-w,y_\lambda)\leqslant\rho\delta+|u|^2+\rho|u|,\end{aligned}$$

由此得出 $|w_\lambda|\leqslant\rho+(|u|^2+\rho|u|)/\delta$，这与(6)一起说明 $y_\lambda,w_\lambda$ 有

界. 若条件(ii)满足,则$(F_\lambda y_\lambda, w_\lambda)\geqslant 0$,

$$\begin{aligned}|F_\lambda y_\lambda|^2 &= (F_\lambda y_\lambda, u-w_\lambda-y_\lambda)\\ &\leqslant (F_\lambda y_\lambda, u-y_\lambda)\leqslant |F_\lambda y_\lambda|\,|u-y_\lambda|,\end{aligned}$$

这与(6)一起推出 $F_\lambda y_\lambda$ 有界.

2° 证 $u\in R(F+G+I)$. 注意由(1)有 $y_\lambda \doteq R_\lambda y_\lambda+\lambda F_\lambda y_\lambda$, $F_\lambda y_\lambda\in FR_\lambda y_\lambda$. 令 $\beta=\sup\limits_{\lambda>0}|F_\lambda y_\lambda|$,对任给 $\lambda,\mu>0$ 有

$$\begin{aligned}|y_\lambda-y_\mu|^2 &= (F_\mu y_\mu-F_\lambda y_\lambda+w_\mu-w_\lambda, y_\lambda-y_\mu)\\ &\leqslant (F_\mu y_\mu-F_\lambda y_\lambda, y_\lambda-y_\mu)\\ &= (F_\mu y_\mu-F_\lambda y_\lambda, R_\lambda y_\lambda-R_\mu y_\mu+\lambda F_\lambda y_\lambda-\mu F_\mu y_\mu)\\ &\leqslant (F_\mu y_\mu-F_\lambda y_\lambda, \lambda F_\lambda y_\lambda-\mu F_\mu y_\mu)\leqslant 2\beta^2(\lambda+\mu),\end{aligned}$$

可见 $y_\lambda-y_\mu\to 0(\lambda,\mu\to 0)$,因此 $y_\lambda\to y(\lambda\downarrow 0)$. 不妨设 $F_\lambda y_\lambda\rightharpoonup v(\lambda\downarrow 0)$,于是 $w_\lambda\rightharpoonup u-v-y(\lambda\downarrow 0)$,从而 $u-v-y\in Gy$(3.1.8). 其次从 $R_\lambda y_\lambda=y_\lambda-\lambda F_\lambda y_\lambda\to y$ 与 $F_\lambda y_\lambda\rightharpoonup v(\lambda\downarrow 0)$推出 $v\in Fy$,因此 $u\in(F+G+I)y$. □

结合 3.4.5、3.3.7 与 3.1.9 得出以下满射结果:

**3.4.6 推论** 设 $F:D\subset X\to 2^{X^*}$ 极大单调,则以下每个条件蕴涵 $R(F+G)=X^*$:(i)$G:C\subset X\to 2^{X^*}$ 极大单调,$(C^\circ\cap D)\cup(C\cap D^\circ)\neq\varnothing$,且 $F+G$ 弱强制;(ii)$G:X\to X^*$ 单调、半连续,$F$ 强单调;(iii)$G:X\to X^*$ 强单调且半连续.

## §5 对积分方程的应用

本节设 $X$ 自反. 给定 $F:X\to X^*$,单调线性算子 $K:X^*\to X$,考虑 Hammerstein 型方程:

$$x+KFx=0. \tag{1}$$

显然,$x\in X$ 满足(1)$\Longleftrightarrow x$ 满足多值算子方程:

$$0\in K^{-1}x+Fx. \tag{2}$$

若 $X$ 为 Hilbert 空间,$K=K^*$,则 $K$ 有唯一"正平方根"$\sqrt{K}$. 若 $z\in X$ 满足方程

$$z + \sqrt{K}F\sqrt{K}z = 0, \tag{3}$$

则 $x = \sqrt{K}z$ 必满足(1). 今将前两节的结果用于方程(2)或(3), 借以得出方程(1)的可解条件. 主要结论汇集于下:

**3.5.1 定理** 以下每个条件保证方程(1)可解: (i) $F$ 单调、有界、半连续且当 $|x|\to\infty$ 时 $\langle Fx,x\rangle>0$ (后一条件蕴涵于 $F$ 强制); (ii) $K$ 强单调, $F$ 单调且半连续; (iii) $K$ 强单调, $F: D\subset X\to X^*$ 极大单调; (iv) $F$ 强单调且半连续; (v) $X$ 为 Hilbert 空间, $K=K^*$, $F$ 单调且半连续; (vi) $X$ 为 Hilbert 空间, $K=K^*$, $\mathrm{Lip}F\leqslant 1/|K|$, $\exists\, c\in[0,1/|K|)$, $\forall\, x\in X$: $(Fx,x)\geqslant -c|x|^2$. 若 $K$, $F$ 单调且其中之一严格单调, 则方程(1)至多有一解.

**证** 首先注意我们已假定 $K$ 单调, 因此 $K$ 连续, 从而 $K$ 与 $K^{-1}$ 皆极大单调(3.1.8, 3.1.9). 若 $K$ 强单调, 则 $K^{-1}$ 单值(3.3.8), 且对任给 $x\in X$ 有

$$\langle K^{-1}x,x\rangle = \langle K^{-1}x, KK^{-1}x\rangle \geqslant \lambda|K^{-1}x|^2 \geqslant \lambda|K|^{-2}|x|^2,$$

其中 $0<\lambda=\text{const}$, 可见 $K^{-1}$ 亦强单调.

当条件(i)满足时直接由 3.3.1 推出(2)可解. 若条件(iii)或(iv)满足, 则用 3.4.6 得出 $R(K^{-1}+F)=X^*$, 从而(2)亦可解. 显然条件(ii)$\Rightarrow$(iii).

以下设 $X$ 为 Hilbert 空间, $K=K^*$. 令 $G=\sqrt{K}F\sqrt{K}$. 若条件(v)满足, 则 $G: X\to X$ 单调且半连续, 因而极大单调, 于是 $R(I+G)=X$(3.3.5), 从而(3)可解. 若条件(vi)满足, 则

$$\mathrm{Lip}G \leqslant |\sqrt{K}|^2\mathrm{Lip}F = |K|\mathrm{Lip}F \leqslant 1,$$

这推出 $I+G$ 单调(见 3.7.3), 从而极大单调. 由

$$\begin{aligned}(x+Gx,x) &= |x|^2 + (F\sqrt{K}x, \sqrt{K}x)\\ &\geqslant |x|^2 - c|\sqrt{K}x|^2 \geqslant (1-c|K|)|x|^2 \quad (\forall\, x\in X)\end{aligned}$$

推出 $I+G$ 强制, 因此亦有 $R(I+G)=X$(3.3.8).

最后的唯一性结论的证明是直接的. □

今考虑将 3.5.1 用于 Hammerstein 积分方程

$$x(t)+\int_\Omega G(t,s)f(s,x(s))\mathrm{d}s, \tag{4}$$

其中$G:\Omega\times\Omega\to L(\mathrm{R}^n)$，$f:\Omega\times\mathrm{R}^n\to\mathrm{R}^n$，$(\Omega,\mu)$是一测度空间，$\mathrm{d}\mu(s)$已记为$\mathrm{d}s$. 令$X=L^p(\Omega,\mathrm{R}^n)$，$1<p<\infty$，则$X$自反且$X^*=L^q(\Omega,\mathrm{R}^n)$，$p^{-1}+q^{-1}=1$. 若定义

$$Kx(t)=\int_\Omega G(t,s)x(s)\mathrm{d}s, x\in X^*; \tag{5}$$

$$Fx(t)=f(t,x(t)), x\in X, \tag{6}$$

则方程(4)转化为算子方程(1)，$K$与$F$的性质分别决定于对$G$与$f$的假设，依次讨论于下.

**3.5.2 命题** 设$G\in L^p(\Omega\times\Omega,L(\mathrm{R}^n))$，则$K:X^*\to X$有定义；$K$单调[强单调]$\Longleftrightarrow$对每个$x\in X^*$有

$$\iint_{\Omega\times\Omega}(G(t,s)x(s),x(t))\mathrm{d}t\mathrm{d}s\geqslant 0[\geqslant\lambda|x|_q^2]$$

$(0<\lambda=\mathrm{const})$；若$G(t,s)^*=G(s,t)$，则$K=K^*$.

**证** 任给$x\in X^*$，由Hölder不等式有

$$\int_\Omega|G(t,s)x(s)|\mathrm{d}s\leqslant|G(t,\cdot)|_p|x|_q. \tag{7}$$

由
$$\int_\Omega|G(t,\cdot)|_p^p\mathrm{d}t=\int_\Omega\mathrm{d}t\int_\Omega|G(t,s)|^p\mathrm{d}s=|G|_p^p<\infty \tag{8}$$

推出$|G(t,\cdot)|_p<\infty$，$\mu$-a.e.. 结合(5)(7)(8)得出$Kx$有定义，且$|Kx|_p^p\leqslant|G|_p^p|x|_q^p$，因此$Kx\in X$. 其余结论是明显的. □

以下总假定$f$满足Caratheodory条件(参考§1.5).

**3.5.3 命题** 若存在$a\in L^q(\Omega,\mathrm{R}_+)$，$\rho\in\mathrm{R}_+$，使得

$$|f(t,x)|\leqslant a(t)+\rho|x|^{p-1}, (t,x)\in\Omega\times\mathrm{R}^n, \tag{9}$$

则$F:X\to X^*$有定义，且有界、半连续.

**证** 任给$x\in X$，$Fx$在$\Omega$上$\mu$-可测；由(9)推出$|Fx|_q\leqslant|a|_q+\rho|x|_p^{p-1}$，因此$Fx\in X^*$且$F:X\to X^*$有界. 其次，$\forall\ x,y,z\in X$，$\lambda\in(0,1)$，有

$|(f(t,x(t)+\lambda y(t)),z(t))|$

$\leqslant\{a(t)+\rho 2^{\rho-1}[|x(t)|^{\rho-1}+|y(t)|^{\rho+1}]\}|z(t)|\in L^1$;

$(f(t,x(t)+\lambda y(t)),z(t))\to(f(t,x(t)),z(t))$, $\mu$-a. e. $(\lambda\downarrow 0)$,

于是可用控制收敛定理得出$\langle F(x+\lambda y),z\rangle\to\langle Fx,z\rangle(\lambda\downarrow 0)$,这表明 $F$ 半连续. □

**注** 在 3.5.3 的条件下实际上可证 $F\in C(X,X^*)$(参考[34; p. 8]),不过本节用不着这一结论.

**3.5.4 命题** 若 $F:X\to X^*$ 有定义,$\forall\, t\in\Omega: f(t,\cdot):\mathrm{R}^n\to\mathrm{R}^n$ 单调[严格单调],则 $F$ 单调[严格单调],且当

$$(f(t,x)-f(t,y),x-y)\geqslant\lambda|x-y|^p(t\in\Omega,x,y\in\mathrm{R}^n)\tag{10}$$

$(0<\lambda=\mathrm{const})$时 $F$ 强单调. 若 $\Omega=J$, $f(\cdot,0)\in L^2(J,\mathrm{R}^n)$, $\forall\, t\in J: f(t,\cdot)$单调, $D=\{x\in L^2(J,\mathrm{R}^n):Fx\in L^2\}$,则 $F:D\subset L^2(J,\mathrm{R}^n)\to L^2(J,\mathrm{R}^n)$极大单调.

**证** 只要证最后一个结论. 令 $X=L^2(J,\mathrm{R}^n)$,仅需证 $R(I+F)=X$(3.3.5). 取定 $y\in X$. $\forall\, t\in J: x\mapsto x+f(t,x)$强单调且连续,故有唯一 $x(t)\in\mathrm{R}^n$ 满足 $x(t)+f(t,x(t))=y(t)$. $\forall\,\varepsilon>0$,取闭集 $J_\varepsilon\subset J$,使 $\mathrm{mes}(J\backslash J_\varepsilon)<\varepsilon$, $f|J_\varepsilon\times\mathrm{R}^n$ 连续[43;Prop. 11.2];不妨设 $y|J_\varepsilon$ 连续. 由

$$|x(t)|^2=(y(t)-f(t,x(t)),x(t))\leqslant(y(t)-f(t,0),x(t))$$

得
$$|x(t)|\leqslant|y(t)|+|f(t,0)|\in L^2(J).\tag{11}$$

由(11)推出$\exists\, r>0: x(J_\varepsilon)\subset B_r(0)$. 因 $f|J_\varepsilon\times B_r(0)$一致连续,故当 $t_k\to t(t_k\in J_\varepsilon)$时 $x(t_k)+f(t,x(t_k))\to y(t)$. 不妨设 $x(t_k)\to z$,于是 $z+f(t,z)=y(t)$,因此 $z=x(t)$, $x(t_k)\to x(t)$,可见 $x|J_\varepsilon$ 连续. 令 $\varepsilon\downarrow 0$ 得出 $x$ 可测;由(11)知 $x\in X$,于是 $y=x+Fx\in R(I+F)$. □

**3.5.5 命题** 设 $F:X\to X^*$ 依(6)有定义;$\exists\, c>0$,使

$$(f(t,x),x)\geqslant c|x|^p,(t,x)\in\Omega\times\mathrm{R}^n,\tag{12}$$

则$\langle Fx,x\rangle\geqslant c|x|_p^p(\forall\, x\in X)$,从而 $F$ 是强制的.

这是明显的.

适当组合 3.5.2~3.5.5 的条件并用 3.5.1,可得出关于方程

(4)的一系列存在性结果.例如,我们有

**3.5.6 定理** 设 $G\in L^p$,依(5)定义的 $K$ 单调,$f$ 满足(9)(12),$\forall\ t\in\Omega$:$f(t,\cdot)$单调,则(4)有 $L^p$ 解.

## §6 对微分方程的应用

考虑 $2m(m\geqslant 1)$阶拟线性边值问题:

$$Lu = f(\text{在 }\Omega\text{ 内}),D^\alpha u|\partial\Omega = 0(|\alpha| \leqslant m-1), \tag{1}$$

其中 $\Omega\subset\mathrm{R}^n$ 是具光滑边界的有界区域,

$$Lu = \sum_{|\alpha|\leqslant m}(-1)^{|\alpha|}D^\alpha A_\alpha(x,Du),Du = (D^\alpha u)_{|\alpha|\leqslant m}, \tag{2}$$

$A_\alpha(\cdot,\cdot):\Omega\times\mathrm{R}^k\to\mathrm{R}^k$,$k$ 是满足 $|\alpha|\leqslant m$ 的 $n$ 重指标 $\alpha\in\mathrm{Z}_+^n$之个数;$f\in L^2(\Omega)$是给定的.令 $A=(A_\alpha)_{|\alpha|\leqslant m}$,定义

$$Nu(x) = A(x,Du(x)); \tag{3}$$

$$a(u,v) = (Nu,Dv); \tag{4}$$

$$b_0(v) = (f,v), \tag{5}$$

其中$(\cdot,\cdot)$记 $L^2$ 内积.令 $X=H_0^m(\Omega)$.若 $u\in X$ 满足

$$a(u,v) = b_0(v)\ (\forall\ v\in X), \tag{6}$$

则称 $u$ 为 BVP(1)的广义解.若有单调算子 $F:X\to X$ 与 $b\in X$,使 $a(u,v)=(Fu,v)_X$,$b_0(v)=(b,v)_X$,$(\cdot,\cdot)_X$ 记 $X$ 中的内积,则(6)转化为算子方程

$$Fu = b. \tag{7}$$

对于(7)的可解性判定,本章已有若干结果可用.上述 $F$ 的存在及其性质取决于对 $A$ 的一列假定:

$(\mathrm{H}_1)$ $A(\cdot,\cdot):\Omega\times\mathrm{R}^k\to\mathrm{R}^k$ 满足 Caratheodory 条件;

$(\mathrm{H}_2)$ $\exists\ g\in L^2(\Omega),\rho>0$:$|A(x,\xi)|\leqslant g(x)+\rho|\xi|$,$(x,\xi)\in\Omega\times\mathrm{R}^k$;

$(\mathrm{H}_3)$ $\forall\ x\in\Omega$:$A(x,\cdot)$单调;

$(\mathrm{H}_4)$ $\exists\ h\in L^1(\Omega),c>0$:$(A(x,\xi),\xi)\geqslant c\sum_{|\alpha|=m}|\xi_\alpha|^2-h(x)$,

$(x,\xi) \in \Omega \times \mathrm{R}^k$;

$(\mathrm{H}_5)$ $\exists\ \beta > 0$: $(A(x,\xi) - A(x,\eta), \xi - \eta) \geqslant \beta \sum_{|\alpha|=m} |\xi_\alpha - \eta_\alpha|^2$, $x \in \Omega, \xi, \eta \in \mathrm{R}^k$.

**3.6.1 定理** 若条件$(\mathrm{H}_1)\sim(\mathrm{H}_4)$满足，则 BVP(1)有解 $u \in X$；若条件$(\mathrm{H}_1)(\mathrm{H}_2)(\mathrm{H}_5)$满足，则(1)有唯一解.

**证** $\forall\, u \in X$，有 $Du \in L^2$；由条件$(\mathrm{H}_1)(\mathrm{H}_2)$及 3.5.3 知 $Nu \in L^2$($N$ 依(3))，且 $N: X \to L^2(\Omega, \mathrm{R}^k)$有界、连续(参考 3.5.3 之后的注). 显然 $a(\cdot,\cdot): X \times X \to \mathrm{R}$ 依(4)有定义，且 $a(u,\cdot)$是线性的. 由 Schwarz 不等式有

$$|a(u,v)| \leqslant |Nu|_2 |Dv|_2 = |Nu|_2 |v|_X,$$

可见 $a(u,\cdot) \in X^*$，于是有唯一 $Fu \in X$ 使 $a(u,v) = (Fu,v)_X (v \in X)$. 这得出算子 $F: X \to X$. 由(5)有

$$|b_0(v)| \leqslant |f|_2 |v|_2 \leqslant |f|_2 |v|_X (\forall\ v \in X),$$

可见 $b_0 \in X^*$，于是有唯一 $b \in X$ 使 $b_0(v) = (b,v)_X$. 因此问题归于证方程(7)在 $X$ 中可解.

$\forall\, u, v \in X$，由条件$(\mathrm{H}_3)$有

$$\begin{aligned}(Fu - Fv, u - v)_X &= (Nu - Nv, Du - Dv) \\ &= \int_\Omega (A(x,Du) - A(x,Dv), Du - Dv)\mathrm{d}x \geqslant 0,\end{aligned}$$

可见 $F$ 单调. 若条件$(\mathrm{H}_5)$满足，则

$$\begin{aligned}(Fu - Fv, u - v)_X &\geqslant \beta \int_\Omega \sum_{|\alpha|=m} |D^\alpha(u-v)|^2 \mathrm{d}x \\ &= \beta \sum_{|\alpha|=m} |D^\alpha(u-v)|_2^2 \geqslant \beta' |u - v|_X^2\end{aligned}$$

($\beta'$是正常数)，最后一个不等号依据 Sobolev 空间 $X$ 的性质. 可见$(\mathrm{H}_5) \Rightarrow F$ 强单调.

$\forall\, u, v, w \in X$，有

$$\begin{aligned}|(Fu - Fv, w)_X| &= |(Nu - Nv, Dw)| \\ &\leqslant |Nu - Nv|_2 |w|_X,\end{aligned}$$

这推出$|Fu-Fv|_X\leqslant|Nu-Nv|_2$. 于是由 $N$ 连续推出 $F$ 连续.

$\forall u\in X$,由条件$(H_4)$有

$$(Fu,u)_X=\int_\Omega(A(x,Du),Du)\mathrm{d}x$$
$$\geqslant c\int_\Omega\sum_{|\alpha|=m}|D^\alpha u|^2\mathrm{d}x-|h|_1\geqslant c'|u|_X^2-|h|_1$$

($c'$为正常数),可见 $F$ 强制.

综上所证,可用 3.3.8 得出定理结论. □

将 3.6.1 用到(1)的特款:

$$-\Delta u+\lambda u=f(\text{在 }\Omega\text{ 内}),u|\partial\Omega=0, \tag{8}$$

其中 $\lambda\geqslant0,f\in L^2(\Omega)$,得到:

**3.6.2 推论** BVP(8)有唯一解 $u\in H_0^1(\Omega)$.

**注** 若以 $H_0^{m,p}(\Omega)$代替 $H_0^m(\Omega)$,则适当修改条件$(H_2)(H_4)(H_5)$后仍可得出类似于 3.6.1 的结果.

其次考虑半线性边值问题:

$$\begin{cases}-Lu+\lambda u=f(x,u),\text{在 }\Omega\text{ 内};\\ D^\alpha u|\partial\Omega=0,|\alpha|\leqslant m-1,\end{cases} \tag{9}$$

其中 $\lambda\in\mathrm{R}$,

$$Lu=\sum_{|\alpha|,|\beta|\leqslant m}(-1)^{|\alpha|}D^\alpha(a_{\alpha\beta}D^\beta u), \tag{10}$$

$a_{\alpha\beta}\in L^\infty(\Omega),f:\Omega\times\mathrm{R}\to\mathrm{R}$. 定义

$$a(u,v)=\sum_{|\alpha|,|\beta|\leqslant m}(a_{\alpha\beta}D^\beta u,D^\alpha v)-\lambda(u,v); \tag{11}$$

$$Fu(x)=f(x,u(x)). \tag{12}$$

如同对问题(1),$u\in X=H_0^m(\Omega)$是(9)的广义解意味着

$$a(u,v)+(Fu,v)=0(\forall v\in X). \tag{13}$$

用 3.6.1 之证可得唯一 $A:X\to X$,使 $a(u,v)=(Au,v)_X(u,v\in X)$,且 $A\in L(X)$. 若有 $N:X\to X$,使$(Fu,v)=(Nu,v)_X(u,v\in X)$,则(13)转化为算子方程

$$Au+Nu=0. \tag{14}$$

在适当条件下,可将 3.4.6 用于(14).

**3.6.3 定理** 设以下条件满足：

($C_1$) $f$ 满足 Caratheodory 条件；

($C_2$) $\exists\ g\in L^2(\Omega),\rho>0:|f(x,u)|\leqslant g(x)+\rho|u|,(x,u)\in\Omega\times\mathrm{R}$；

($C_3$) $\forall\ x\in\Omega:f(x,\cdot)$单调增；

($C_4$) $L$ 是强椭圆的，即$\exists\ c_0>0$，使得

$$\sum_{|\alpha|,|\beta|=m} a_{\alpha\beta}(x)\xi_\alpha\xi_\beta \geqslant c_0\sum_{|\alpha|=m}|\xi_\alpha|^2, x\in\Omega,(\xi_\alpha)\in\mathrm{R}^k,$$

则当 $\lambda$ 适当小时 BVP(9)有唯一解 $u\in X$.

**证** 结合条件($C_1$)～($C_3$)与 3.5.3、3.5.4 得出，(12)定义一个连续单调算子 $F:L^2(\Omega)\to L^2(\Omega)$. $\forall\ u\in X$，由

$$|(Fu,v)|\leqslant|Fu|_2|v|_2\leqslant|Fu|_2|v|_X(\forall\ v\in X)$$

知 $Fu\in X^*$，于是有唯一 $Nu\in X:(Nu,v)_X=(Fu,v)$. 由

$$|Nu-Nv|_X\leqslant|Fu-Fv|_2$$

与
$$(Nu-Nv,u-v)_X=(Fu-Fv,u-v)$$

推出 $N:X\to X$ 亦是连续单调算子.

因 $L$ 是强椭圆的，由 Gårding 不等式有

$$(Au,u)_X=a(u,u)\geqslant c|u|_X^2-(C+\lambda)|u|_2^2(\forall\ u\in X),\tag{15}$$

$c>0,C\in\mathrm{R}$ 是常数. 可见当 $\lambda\leqslant -C$ 时 $A$ 强单调.

综上所证，可用 3.4.6 得出所要结论. □

对于(9)的特款：

$$\Delta u=f(x,u)(在\ \Omega\ 内),u|\partial\Omega=0,\tag{16}$$

由 Poincaré 不等式知(15)中可取 $C=0$，因此有

**3.6.4 推论** 若 3.6.3 中条件($C_1$)～($C_3$)满足，则 BVP(16)有唯一解 $u\in H_0^1(\Omega)$.

下面循另一思路来解 BVP(9). 令 $Y=L^2(\Omega)$，已知条件($C_1$)～($C_3$)推出 $F:Y\to Y$ 为连续单调算子，$F$ 依(12). 令 $Au=Lu-\lambda u$，假定对适当选定的 $D(A)\subset Y,A:D(A)\to Y$ 有定义，且 $D(A)\subset$

$H_0^m(\Omega)$. 于是 BVP(9)转化为算子方程

$$Au + Fu = 0. \tag{17}$$

仍可由条件($C_4$)推出 $A$ 强单调(假定 $\lambda$ 适当小). 由 3.4.6,只需验证 $A$ 极大单调,即可断定(17)有解. 由 3.3.5,这归于证 $\exists\ \mu>0$: $R(A+\mu I)=Y$,即 BVP

$$\begin{cases} Lu + (\mu - \lambda)u = \varphi \in Y(\text{在}\ \Omega\ \text{内}) \\ D^\alpha u \mid \partial\Omega = 0, |\alpha| \leqslant m - 1 \end{cases}$$

恒有解 $u \in D(A)$. 这一结论可从适当的线性 BVP 结果得出.

作为说明,考虑简单的 1 维问题

$$x'' = f(t,x)(0 < t < 1), x(0) = x(1) = 0. \tag{18}$$

令 $Y=L^2(J)$, $Ax=-x''$, $D(A)=\{x\in C^1(J): x''\in Y, x(0)=x(1)=0\}$. 由分部积分与 Wirtiager 不等式[63;p. 206]有

$$(Ax,x) = |x'|_2^2 \geqslant \pi^2 |x|_2^2 (\forall\ x \in D(A)),$$

可见 $A: D(A)\subset Y\to Y$ 强单调. $\forall\ y\in Y$,直接验知

$$x'' - x = y(t)(0 < t < 1), x(0) = x(1) = 0$$

恒有解 $x\in D(A)$,因此 $A$ 极大单调. 于是当 $f$ 满足条件($C_1$)~($C_3$)时 BVP 有唯一 $L^2$ 解.

## §7 增 生 算 子

本节设 $F: D\subset X\to 2^X$. 若 $X$ 是 Hilbert 空间,则 $F$ 单调$\Longleftrightarrow$ $\forall\ x,y\in D: (Fx-Fy,x-y)\geqslant 0$. 若以半内积代替内积,则在 Banach 空间的框架内另辟一引入单调性的途径,所得结果与前此所述的单调算子理论略相类似且各有短长,因而可互相补充.

**3.7.1定义** 若$\forall\ x,y\in D: x\neq y\Rightarrow (Fx-Fy,x-y)_+\geqslant 0$ [$>0$],则称 $F$ 为增生[严格增生]算子. 若$\exists\ \lambda>0, \forall\ x,y\in D: (Fx-Fy,x-y)_+\geqslant\lambda|x-y|^2$,则称 $F$ 为强增生(或 $\lambda$-强增生)算子. 若 $F$ 增生且$(Fx-v,x-y)_+\geqslant 0(\forall\ x\in D)\Rightarrow v\in Fy$,则称 $F$ 为极大增生算子;若 $F$ 增生且$\exists\ \lambda>0: R(I+\lambda F)=X$,则称 $F$ 为 m-增生

算子.

如§2中所见,$(\cdot,\cdot)_+$的性质对空间的"凸性"依赖颇大.增生算子亦然,例如,易直接证明.

**3.7.2 命题** 设$F$极大增生.若$X^*$严格凸,则$\forall x\in D$:$Fx$为凸闭集;若$X$自反,$X^*$局部一致凸,则$F$是$sw$-闭的(参考3.1.1,3.1.8,3.2.5).

$\forall \lambda>0$,称$R_\lambda\triangleq(I+\lambda F)^{-1}$为预解式,称$F_\lambda\triangleq\lambda^{-1}(I-R_\lambda)$为$F$的Yosida逼近,它们有对应于§4(1)的关系:

$$I-R_\lambda=\lambda F_\lambda, F_\lambda\subset FR_\lambda. \tag{1}$$

若$X$自反,$X$与$X^*$严格凸,$F$极大增生,则有$F$的唯一选择$F_0:D\to X$满足(参考§4(2),3.7.2,4.1.10)

$$F_0x\in Fx, |F_0x|=d(0,Fx),\ x\in D. \tag{2}$$

为展开平行于3.4.2的结果,首先建立

**3.7.3 命题** (i)$F$是增生的$\Longleftrightarrow\forall(x_i,y_i)\in \mathrm{Gr}F(i=1,2),\lambda>0$:$|x_1-x_2+\lambda(y_1-y_2)|\geqslant|x_1-x_2|$;因此若$F$单值,$\mathrm{Lip}F\leqslant1$,则$I\pm F$是增生的.(ii)$F$是增生的$\Longleftrightarrow\forall\lambda>0$:$R_\lambda$单值且$\mathrm{Lip}R_\lambda\leqslant1$(从而$F_\lambda$增生且$\mathrm{Lip}F_\lambda\leqslant2/\lambda$).(iii)若$F$是m-增生的,则$F$极大增生,且$\forall\lambda>0$:$R(I+\lambda F)=X$,因而$R_\lambda:X\to D$与$F_\lambda:X\to X$有定义.

**证** 结合3.7.1与§2(1)直接得出(i).

(ii)若$F$是增生的,$\lambda>0,x_i\in R_\lambda y_i(i=1,2)$,则$y_i-x_i\in\lambda Fx_i$,于是(参看3.2.2(vii))

$0\leqslant(y_1-x_1-y_2+x_2,x_1-x_2)_+=(y_1-y_2,x_1-x_2)_+-|x_1-x_2|^2$,

因此$|x_1-x_2|\leqslant|y_1-y_2|$,这表明$R_\lambda$单值且$\mathrm{Lip}R_\lambda\leqslant1$.反之,若$R_\lambda$单值且$\mathrm{Lip}R_\lambda\leqslant1(\forall\lambda>0)$,$y_i\in Fx_i(i=1,2)$,则$x_i=R_\lambda(x_i+\lambda y_i)$,$|x_1-x_2|\leqslant|x_1-x_2+\lambda(y_1-y_2)|$,可见$F$是增生的.

(iii)设$\mu>0$:$R(I+\mu F)=X$.若$\lambda>\mu/2,y\in X$,定义$Gx=R_\mu((1-\rho)x+\rho y),\rho=\mu/\lambda$,则$\mathrm{Lip}G\leqslant|1-\rho|<1$,于是$\exists x\in X$:

$Gx=x, y\in(I+\lambda F)x$. 由已证结论可推出：$A\triangleq\{\lambda>0: R(I+\lambda F)=X\}$在$(0,\infty)$中既开且闭，因此 $A=(0,\infty)$. 若$(Fx-v, x-y)_+\geqslant 0(\forall\, x\in D)$，则取 $x=R_1(y+v)$得$(y-x, x-y)_+\geqslant 0$，这得出 $x=y, v\in Fy$，因此 $F$ 为极大增生. □

**3.7.4 定理** 设 $F$ 是 m-增生的，$\lambda>0$. 则有：(i)$\forall\, x\in\overline{D}: R_\lambda x\to x(\lambda\downarrow 0)$；(ii)$\forall\, x\in D: |F_\lambda x|\leqslant \mathrm{d}(0,Fx)$；$\forall\, x\in X\backslash\overline{D}: |F_\lambda x|\to\infty(\lambda\downarrow 0)$；$\forall\, x\in X: \lambda\mapsto|F_\lambda x|$单调减. (iii)若 $X$ 自反、严格凸，$X^*$局部一致凸，$x\in D$，则 $F_\lambda x\rightharpoonup F_0x$，$|F_\lambda x|\to|F_0x|(\lambda\downarrow 0)$；若 $X$ 亦局部一致凸，则 $F_\lambda x\to F_0x(\lambda\downarrow 0)$. (iv)若 $X$ 局部一致凸，则 $\overline{D}$ 凸.

**证** 证法与 3.4.2 略相类似. $\forall\, x\in X, y\in D, u\in Fy$，由 $y=R_\lambda(y+\lambda u)$与 $\mathrm{Lip}R_\lambda\leqslant 1$ 得

$$|y-R_\lambda x|\leqslant|y-x|+\lambda|u|. \tag{3}$$

在(3)中令 $y=x\in D$ 得 $\lambda|F_\lambda x|=|x-R_\lambda x|\leqslant\lambda\mathrm{d}(0,Fx)$，因此$|F_\lambda x|\leqslant\mathrm{d}(0,Fx)$，$R_\lambda x\to x(\lambda\downarrow 0)$. 由于 $\mathrm{Lip}R_\lambda\leqslant 1$，对 $x\in\overline{D}$ 亦有 $R_\lambda x\to x(\lambda\downarrow 0)$. 当 $x\in X\backslash\overline{D}$ 时显然有$|F_\lambda x|\to\infty(\lambda\downarrow 0)$. 若 $x\in X, \lambda>\mu>0$，则

$$\begin{aligned}\lambda|F_\lambda x|-\mu|F_\mu x|&\leqslant|\lambda F_\lambda x-\mu F_\mu x|=|R_\lambda x-R_\mu x|\\&=|R_\mu(I+\mu F)R_\lambda x-R_\mu x|=|R_\mu(R_\lambda x+\mu F_\lambda x)-R_\mu x|\\&\leqslant|R_\lambda x+\mu F_\lambda x-x|=(\lambda-\mu)|F_\lambda x|,\end{aligned}$$

由此显然推出$|F_\lambda x|\leqslant|F_\mu x|$. 这就证得(i)(ii).

证(iii). 取定 $x\in D$. 由(ii)有$|F_\lambda x|\leqslant|F_0x|$. 任给 $\lambda_n\downarrow 0$，不妨设 $F_{\lambda_n}x\rightharpoonup u$. 因 $F_\lambda x\in FR_\lambda x, R_{\lambda_n}x\to x$，而 $F$ 为 sw-闭(3.7.2)，故 $u\in Fx$. 由

$$|u|\leqslant\varliminf_n|F_{\lambda_n}x|\leqslant\varlimsup_n|F_{\lambda_n}x|\leqslant|F_0x|\leqslant|u|$$

得 $u=F_0x$，$|F_{\lambda_n}x|\to|F_0x|$，$F_{\lambda_n}x\rightharpoonup F_0x$. 这证得 $F_\lambda x\rightharpoonup F_0x$，$|F_\lambda x|\to|F_0x|(\lambda\downarrow 0)$. 若 $X$ 局部一致凸，则 $F_\lambda x\to F_0x(\lambda\downarrow 0)$.

证(iv). 只需证$\forall\, x,y\in D, t\in J, z=(1-t)x+ty, \lambda_n\downarrow 0$，有$R_{\lambda_n}z\to z$. 由(3)知$R_{\lambda_n}z$有界，不妨设$R_{\lambda_n}z\rightharpoonup u$. 由(3)推出$|x-u|\leqslant$

$|x-z|$，$|y-u|\leqslant|y-z|$，这结合

$$|x-z|+|y-z|=|x-y|\leqslant|x-u|+|y-u|$$

得出 $|x-u|=|x-z|$，$|y-u|=|y-z|$；于是由 $X$ 严格凸得出 $z=u$，因此 $R_{\lambda_n}z\rightharpoonup z$，从而 $R_{\lambda_n}z-x\rightharpoonup z-x$. 由(3)有

$$\overline{\lim_n}|R_{\lambda_n}z-x|\leqslant|x-z|,$$

于是 $|R_{\lambda_n}z-x|\to|z-x|$，因此 $R_{\lambda_n}z-x\to z-x$，从而 $R_{\lambda_n}z\to z$. □

## §8 非扩张半群

设 $F:D\subset X\to X$，考虑初值问题(IVP)：

$$x'=-Fx,x(0)=\xi\in D. \tag{1}$$

若 $x\in C([0,b),D)$，$b>0$，弱导数 $x'$ 存在(这意味着 $\tau\to0\Rightarrow\tau^{-1}\Delta x(t,\tau)\rightharpoonup x'(t)$)且(1)满足，则称 $x$ 为IVP(1)在 $[0,b)$ 上的解.

**3.8.1 引理** 设 $x:(a,b)\to X$ 可微，或 $x$ 弱可微而 $X^*$ 严格凸，$\varphi(t)\coloneqq|x(t)|$，则 $D^-\varphi^2(t)=2\varphi(t)D^-\varphi(t)=2(x'(t),x(t))_-$，此处 $D^-\varphi(t)\triangleq\overline{\lim_{\lambda\uparrow0}}\lambda^{-1}\Delta\varphi(t,\lambda)$ 是"左上导数".

**证** 不妨只考虑 $x$ 弱可微 $X^*$ 严格凸的情况. 取定 $t\in(a,b)$，可设 $z\triangleq x(t)\neq0$. 令 $\alpha=\lambda^{-1}\Delta x(t,\lambda)-x'(t)$，则 $\alpha\rightharpoonup o(\lambda\to0)$. 因 $X^*$ 严格凸，故有 $|z+\lambda y|-|z|=\langle|z|^{-1}\mathscr{D}z,\lambda y\rangle+o(\lambda)(\lambda\to0,y\in X)$，$\mathscr{D}$ 记对偶映射(3.2.5(ii)). 于是

$$\Delta\varphi(t,\lambda)=|z+\lambda x'(t)+\lambda\alpha|-|z|=\frac{\lambda}{|z|}\langle\mathscr{D}z,x'(t)+\alpha\rangle$$
$$+o(\lambda),$$

这推出 $\varphi(t)D^-\varphi(t)=\langle\mathscr{D}z,x'(t)\rangle=(x'(t),x(t))_-$. □

**3.8.2 引理** 设 $\omega\in C([0,b],\mathrm{R})$ 满足 $D^-\omega(t)\leqslant\alpha\omega(t)+\beta$，$\alpha$，$\beta\in\mathrm{R}$，则当 $\alpha=0$ 时 $\omega(t)\leqslant\omega(0)+\beta t$；当 $\alpha\neq0$ 时

$$\omega(t)\leqslant\omega(0)e^{\alpha t}+\frac{\beta}{\alpha}(e^{\alpha t}-1),0\leqslant t\leqslant b. \tag{2}$$

这是一个标准的微分不等式结论(如参考[65]).

下面在两组不同的条件下证明(1)的可解性.

**3.8.3 定理** 设存在 $k\in\mathrm{R}$,使得

$$(Fx-Fy,x-y)_+\geqslant k|x-y|^2,x,y\in D;\qquad (3)$$

$F$ 连续,或 $F$ 次连续而 $X^*$ 一致凸,则对任给 $\xi\in D^\circ$,IVP(1)有解;当 $D=X$ 时(1)在 $\mathrm{R}_+$ 上有唯一解.

**证** 只考虑 $F$ 次连续 $X^*$ 一致凸的情况,$F$ 连续的情况可类似证明. 由 $F$ 次连续推出 $F$ 局部有界,故有 $\delta,r>0$:$F\overline{B}_\delta(\xi)\subset B_r(0)$. 令 $b=\delta/r$. 定义"Peano 逼近序列"$\{x_n\}$如下:当 $t<0$ 时令 $x_n(t)=0$;

$$x_n(t)=\xi-\int_0^t Fx_n\left(s-\frac{1}{n}\right)\mathrm{d}s,0\leqslant t\leqslant b,n\geqslant 1,\qquad (4)$$

(4)中的积分是 Pettis 积分. 记 $y_n(t)=x_n(t-n^{-1})$. 显然 $x_n(0)=\xi$, $|x_n(t)-x_n(s)|\leqslant r|t-s|,t,s\in[0,b]$;$x_n$ 弱可微且 $x'_n(t)=-Fy_n(t),t\in[0,b]$. 取定 $m,n\geqslant 1$,令 $\varphi(t)=|x_m(t)-x_n(t)|$,则由 3.8.1 与条件(3)有(下式中省去了变量 $t\in[0,b]$):

$$\begin{aligned}D^-\varphi^2&=2(x'_m-x'_n,x_m-x_n)_-=-2(Fy_m-Fy_n,x_m-x_n)_+\\&=-2(Fy_m-Fy_n,y_m-y_n)_++\alpha_{mn}\leqslant-2k|y_m-y_n|^2+\alpha_{mn}\\&\leqslant 2|k|\left(\varphi+\frac{r}{m}+\frac{r}{n}\right)^2+\alpha_{mn}\\&\leqslant 4|k|\left[\varphi^2+\left(\frac{r}{m}+\frac{r}{n}\right)^2\right]+\alpha_{mn}\leqslant 4|k|\varphi^2+\beta_{mn},\end{aligned}$$

其中 $\alpha_{mn}=2(Fy_m-Fy_n,y_m-y_n)_--2(Fy_m-Fy_n,x_m-x_n)$;

$$\beta_{mn}=4|k|\left(\frac{r}{m}+\frac{r}{n}\right)^2+\sup_{0\leqslant t\leqslant b}\alpha_{mn}(t).$$

由 $(\cdot,\cdot)_+$ 在有界集上一致连续(3.2.5(vi))得出 $\alpha_{mn}(t)\rightrightarrows 0$,从而 $\beta_{mn}\to 0(m,n\to\infty)$. 对 $\omega=\varphi^2$ 应用 3.8.2 并注意 $\varphi(0)=0$ 得 $\varphi^2(t)\leqslant(\beta_{mn}/4|k|)(e^{4|k|t}-1)$(当 $k=0$ 时 $\varphi^2(t)\leqslant\beta_{mn}t$),因此 $x_m(t)-x_n(t)\rightrightarrows 0(m,n\to\infty,t\in[0,b])$. 于是有 $x\in C([0,b],D)$:$x_n(t)\rightrightarrows x(t)$;由(4)易直接验知 $x$ 是 IVP(1)的解.

若 $x,z$ 是(1)在$[0,b]$上的解，则由 3.8.1 与(3)得出 $\psi(t)\triangleq|x(t)-z(t)|$满足$D^-\psi^2(t)\leqslant-2k\psi^2(t)$，于是由3.8.2有$\psi(t)\leqslant\psi(0)e^{-kt}=0$，这得出 $x(t)\equiv z(t)$.

若(1)的解 $x$ 的极大正半定义区间为$[0,\rho]$，$0<t<t+\lambda<\rho$，则类似于上面的论证对$|x(t+\lambda)-x(t)|$用3.8.1与3.8.2得出$|x(t+\lambda)-x(t)|\leqslant|x(\lambda)-x(0)|e^{-kt}$. 由此推出，当 $\rho<\infty$时 $x(\rho-0)$必存在，从而当 $D=X$ 时 $x(t)$可延拓到 $t=\rho$ 之右侧，得出矛盾. 因此 $\rho=\infty$. 用上段所得之唯一性结论易推出 $x$ 是(1)在 $\mathrm{R}_+$ 上的唯一解. □

**3.8.4 定理** 设 $X^*$ 一致凸，$F$ 是 m-增生的，则 IVP(1)在 $\mathrm{R}_+$ 上有唯一解 $x$ 且$|x'(t)|$单调减.

**证** $\forall\ \lambda>0$，由 $\mathrm{Lip}F_\lambda\leqslant2/\lambda$(3.7.3)知 IVP“$x'=-F_\lambda x,x(0)=\xi\in D$”在$\mathrm{R}_+$上有唯一$C^1$解$x_\lambda$. $\forall\ \tau>0$，对$\varphi(t)\triangleq|x_\lambda(t+\tau)-x_\lambda(t)|$用3.8.1并注意$F_\lambda$增生得$D^-\varphi(t)\leqslant0$；于是由3.8.2知$\varphi(t)$单调减，从而$|x'_\lambda(t)|$单调减，特别有

$$|F_\lambda x_\lambda(t)|=|x'_\lambda(t)|\leqslant|x'_\lambda(0)|=|F_\lambda\xi|\leqslant|F\xi|,t\geqslant0. \tag{5}$$

$\forall\ \lambda,\mu>0$，对 $\psi(t)\triangleq|x_\lambda(t)-x_\mu(t)|\ (t\geqslant0)$用 3.8.1 得

$$\begin{aligned}\psi D^-\psi&=(FR_\mu x_\mu-FR_\lambda x_\lambda,x_\lambda-x_\mu)_+\\&\leqslant(FR_\lambda x_\lambda-FR_\mu x_\mu,R_\lambda x_\lambda-R_\mu x_\mu)_+\\&\quad-(FR_\lambda x_\lambda-FR_\mu x_\mu,x_\lambda-x_\mu)_+\triangleq a_{\lambda\mu},\end{aligned} \tag{6}$$

其中已省去变元 $t$. 利用(5)推出

$$|R_\lambda x_\lambda(t)-x_\lambda(t)|=\lambda|F_\lambda x_\lambda(t)|\leqslant\lambda|F\xi|\to0(\lambda\downarrow0).$$

$\forall\ b>0$，用$(\cdot,\cdot)_+$在有界集上一致连续推出 $a_{\lambda\mu}(t)\rightrightarrows0(\lambda,\mu\downarrow0,t\in[0,b])$，于是结合(6)与 3.8.2 得$|x_\lambda(t)-x_\mu(t)|\rightrightarrows0(\lambda,\mu\downarrow0,t\in[0,b])$. 因此存在 $x\in C(R_+,X)$，$\forall\ b>0$：$x_\lambda(t)\rightrightarrows x(t)\ (\lambda\downarrow0,t\in[0,b])$. 显然 $x(0)=\xi$，$R_\lambda x_\lambda(t)\to x(t)\ (\lambda\downarrow0)$. 取定 $t\geqslant0$，任给 $\lambda_n\downarrow0$，不妨设 $F_{\lambda_n}x_{\lambda_n}(t)\rightharpoonup y$，这与 $R_{\lambda_n}x_{\lambda_n}(t)\to x(t)$及 3.7.2 一起推出 $y=Fx(t)$，因此有$F_\lambda x_\lambda(t)\rightharpoonup Fx(t)$. 用此易推出$Fx(t)$弱连续，且

$x(t)=\xi-\int_0^t Fx(s)\mathrm{d}s\ (t\geqslant 0)$，这表明 $x$ 是(1)在 $\mathrm{R}_+$ 上的解．唯一性的证明如同 3.8.3. □

设 3.8.3 或 3.8.4 之条件满足(对前者假定 $D=X$)，以 $u(\cdot,\xi)$ 记 IVP(1)的解，令 $T_t=u(t,\cdot)$，则得到确定的映射 $u:\mathrm{R}_+\times D\to D$ 与算子族 $\{T_t:D\to D\mid t\geqslant 0\}$，后者有性质：(i) $T_0=I$，$T_tx\to x$，$t^{-1}(T_tx-x)\rightharpoonup -Fx(t\downarrow 0,x\in D)$；(ii) $T_{t+s}=T_tT_s$；(iii) $\mathrm{Lip}T_t\leqslant e^{-kt}$，$k$ 依(3)，在 3.8.4 之条件下取 $k=0$. 性质(i)(ii)是明显的．其次，对 $|T_tx-T_ty|$ 用条件(3)与3.8.1，3.8.2得出 $|T_tx-T_ty|\leqslant |x-y|e^{-kt}$(在 3.8.4 条件下取 $k=0$)，可见 $\mathrm{Lip}T_t\leqslant e^{-kt}$. 称如上的 $T_t$ 为 $D$ 上由 $-F$ 生成的(非线性)算子半群．若 $k=0$(当 $F$ 增生时如此)，则 $\mathrm{Lip}T_t\leqslant 1$，即 $T_t$ 是非扩张的，此时称 $T_t$ 为非扩张半群.

依上面的记号，若 $T_tx\equiv x(t\geqslant 0)$，则必 $Fx=0$. 用此可得出一些满射性结果.

**3.8.5 定理** 以下每个条件蕴涵 $R(F)=X$：(i) $F:X\to X$ 强增生、连续(或次连续而 $X^*$ 一致凸)；(ii) $X^*$ 一致凸，$F:X\to X$ 增生、强连续(即 $x_n\rightharpoonup x\Rightarrow Fx_n\to Fx$)且弱强制；(iii) $X$ 与 $X^*$ 一致凸，$F:D\to X$ 为 m-增生且弱强制.

**证** 只要证 $0\in R(F)$. 以 $T_t$ 记由 $-F$ 生成的半群.

若条件(i)满足，则有 $k>0$：$\mathrm{Lip}T_1\leqslant e^{-k}<1$，于是有 $\xi\in\mathrm{Fix}T_1$，从而 $x(t)\triangleq T_t\xi$ 是 IVP(1)的 1-周期解．$\forall\, t>0$，由周期性有

$$|x(t)-x(0)|=|T_1x(t)-T_1x(0)|\leqslant e^{-k}|x(t)-x(0)|,$$

这推出 $x(t)\equiv x(0)(t\geqslant 0)$，因此 $F\xi=0$.

若条件(ii)满足，则对 $F+n^{-1}I$ 用已证结论得 $x_n\in X$：$Fx_n+n^{-1}x_n=0(n\geqslant 1)$. 由

$$|Fx_n||x_n|=(-Fx_n,x_n)_+\leqslant(-F0,x_n)_+\leqslant|F0||x_n|$$

推出 $|Fx_n|\leqslant|F0|$，从而 $x_n$ 有界，因此 $Fx_n\to 0$，且可设 $x_n\rightharpoonup x$，这得出 $Fx=0$.

若条件(iii)满足，则由 $\mathrm{Lip}T_t\leqslant 1$ 推出 $(t,x)\mapsto T_tx$ 可扩张到

$R_+\times\overline{D}$ 上，使 $T_t$ 为 $\overline{D}$ 上的半群，且保持 $\mathrm{Lip}(T_t|\overline{D})\leqslant 1$. 取定 $\xi\in D$，令 $x(t)=T_t\xi$，则由3.8.4有 $|Fx(t)|=|x'(t)|\leqslant|x'(0)|=|F\xi|$，因此 $x(R_+)$有界. 由 $X$ 一致凸推出 $\overline{D}$ 凸(3.7.4). 利用下面引述的不动点结果(3.8.6)，得 $x_0\in\overline{D}:T_tx_0\equiv x_0$. $\forall\ x\in D,t>0$，有

$$(x-T_tx,x-x_0)_+=|x-x_0|^2+(T_tx_0-T_tx,x-x_0)_+$$
$$\geqslant|x-x_0|^2-(\mathrm{Lip}T_t)|x-x_0|^2\geqslant 0;$$

$$(Fx,x-x_0)_+=\lim_{t\downarrow 0}\frac{1}{t}(x-T_tx,x-x_0)_+\geqslant 0.$$

由 $F$ 极大增生得 $Fx_0=0$. □

证明中用到以下结果(参考[43;17.8])：

**3.8.6 定理** 若 $X$ 一致凸，$D\subset X$ 凸闭，$T_t$ 是 $D$ 上的非扩张半群，$\exists\ \xi\in D:\{T_t\xi:t\geqslant 0\}$有界，则$\exists\ x_0\in D:T_tx_0\equiv x_0$.

由 3.8.5 推出 3.1.9(i)与 3.4.2(iv)的以下类似：

**3.8.7 推论** (i)若 $X^*$ 一致凸，$F:X\to X$ 增生、次连续，则 $F$ 是 m-增生的. (ii)若 $F:D\subset X\to 2^X$ 是 m-增生的，则$\forall\ \lambda>0:F_\lambda$ 是 m-增生的.

**证** (i)因 $F+I$ 是强增生与次连续的，故由 3.8.5 有 $R(F+I)=X$，从而 $F$ 为 m-增生.

(ii)因 $F_\lambda+I(\lambda>0)$是强增生、连续的(3.7.3)，故由 3.8.5 有 $R(F_\lambda+I)=X$，从而 $F_\lambda$ 为 m-增生. □

§4 中的和定理对于增生算子亦有适当的推广，不拟详述. 下面仅给出 3.4.4 的以下类似.

**3.8.8 命题** 设 $X^*$ 严格凸，$F:D\subset X\to 2^X$ 与 $G:C\subset X\to 2^X$ 是 m-增生的，$\lambda>0$，则 $F_\lambda+G$ 是 m-增生的.

**证** 直接验知 $F_\lambda+G$ 是增生的($X^*$ 严格凸用于此!). 取定 $\lambda>0$；取 $\mu>2/\lambda,y\in X$，定义

$$Tx=(G+\mu I)^{-1}(y-F_\lambda x),x\in X,$$

则
$$\mathrm{Lip}T\leqslant\mathrm{Lip}(G+\mu I)^{-1}\mathrm{Lip}F_\lambda\leqslant\frac{2}{\lambda\mu}<1$$

(用3.7.3)，于是有$x\in\mathrm{Fix}T$. 直接验知$y\in F_\lambda x+Gx+\mu x$，因此

$R(F_\lambda+G+\mu I)=X$，从而 $F_\lambda+G$ 为 m-增生. □

## §9 逼近解

给定单调算子 $F:X\to X^*$ 与 $b\in X^*$，本章解答了关于算子方程

$$Fx=b \tag{1}$$

的如下基本问题：在什么条件下方程(1)有解、有唯一解且解连续依赖于 $b$? 至于解的计算，则有多种逼近方法，它们都是从一定的实际算法提炼出来的，本节讨论的 Ritz 投影法(3.9.1)与广义梯度法(3.9.2)是其典型例子.

**3.9.1 定理**(Bowder-Minty，1963) 设 $X$ 是具可数基$\{e_n\}$的自反空间，$X_n=\text{span}\{e_1,\cdots,e_n\}$；$F$ 单调、半连续且强制，则以下结论成立：(i)方程(1)的解构成非空有界闭凸集；(ii) Galerkin 方程

$$\langle Fx-b,e_i\rangle=0\ (1\leqslant i\leqslant n) \tag{2}$$

有解 $x_n\in X_n$，$\{x_n\}$有子列弱收敛于(1)的某个解；(iii)若 $F$ 严格单调，则(1)与(2)分别有唯一解 $x$ 与 $x_n\in X_n$，且 $x_n\rightharpoonup x$；当 $F$ 强单调时 $x_n\to x$.

**证** 在定理条件下 $F$ 与 $F^{-1}$皆极大单调(3.1.9)，于是直接由 3.1.8 与 3.3.8 得出(i).

(ii)将 3.3.8 用到映射 $X_n\to X_n^*$，$x\mapsto Fx|X_n$ 得出方程(2)有解 $x_n\in X_n$. 显然 $Fx_n-b\in X_n^\perp$. 由

$$\langle Fx_n,x_n\rangle=\langle b,x_n\rangle\leqslant|b||x_n| \tag{3}$$

及 $F$ 强制得出 $x_n$ 有界，因此$\{x_n\}$含弱收敛子列. 为记号简便，不妨设 $x_n\rightharpoonup x$，今证 $Fx=b$. 由 $F$ 极大单调，只需证$\langle Fy-b,y-x\rangle\geqslant0$ $(\forall\ y\in X)$. 取定 $y\in X$，因 $X_0\triangleq\bigcup_1^\infty X_n$ 满足 $\overline{X}_0=X$，故有 $y_k\in X_0$：$y_k\to y$. 由 $F$ 次连续有 $Fy_k\rightharpoonup Fy$. 固定 $k$，当 $n$ 充分大时有$\langle Fx_n,y_k-x_n\rangle=\langle b,y_k-x_n\rangle$，于是

$$\langle Fy_k-b,y_k-x_n\rangle=\langle Fy_k-Fx_n,y_k-x_n\rangle\geqslant0.$$

先令 $n\to\infty$，次令 $k\to\infty$，得出 $\langle Fy-b,y-x\rangle\geqslant 0$.

(iii) 当 $F$ 严格单调时显然(1)与(2)分别有唯一解 $x$ 与 $x_n\in X_n$. 由(ii)之证看出，$\{x_n\}$ 的每个子列有子列弱收敛于 $x$，因此 $x_n\rightharpoonup x$. 取 $r,\rho>0$，使得 $|y|\leqslant r\Rightarrow|Fy|\leqslant\rho$(3.1.4). $\forall\ y\in X$，由 $F$ 单调有 $\langle Fx_n,y\rangle\leqslant\langle Fx_n,x_n\rangle-\langle Fy,x_n-y\rangle$；这结合(3)推出

$$\begin{aligned}|Fx_n|&=\sup_{|y|=r}\frac{1}{r}\langle Fx_n,y\rangle\\&\leqslant\frac{1}{r}(|b||x_n|+\rho|x_n|+\rho r),\end{aligned}$$

这表明 $Fx_n$ 有界. $\forall\ y\in X_0$，当 $n$ 充分大时有 $\langle Fx_n,y\rangle=\langle b,y\rangle$；这一事实结合 $Fx_n$ 有界得出 $Fx_n\rightharpoonup b$. 由(3)与 $x_n\rightharpoonup x$ 得出 $\langle Fx_n,x_n\rangle\to\langle b,x\rangle$. 若 $F$ 为 $\lambda$-强单调，则结合以上结论有

$$\begin{aligned}\lambda|x_n-x|^2&\leqslant\langle Fx_n-b,x_n-x\rangle\\&=\langle Fx_n,x_n\rangle-\langle Fx_n-b,x\rangle-\langle b,x_n\rangle\to 0,\end{aligned}$$

因此 $x_n\to x(n\to\infty)$. □

另一类逼近方法基于以下考虑：(1)等价于

$$x=x-tA(Fx-b)\triangleq Gx, \tag{4}$$

其中 $t>0$，$A:X^*\to X$ 是满足 $Au=0\Longleftrightarrow u=0$ 的某个算子. 于是问题是，应如何限定 $F$ 与选定 $t,A$，使得如下定义的迭代序列收敛于 $G$ 的不动点：

$$x_{k+1}=x_k-tA(Fx_k-b)(k\geqslant 0);x_0=0. \tag{5}$$

迭代方法通常要求一定的 Lipschitz 条件.

**3.9.2 定理** 设 $X$ 是 Hilbert 空间；$F:X\to X$ 为 $\lambda$-强单调，$\forall\ r>0$：$L(r)\triangleq\mathrm{Lip}(F|\overline{B}_r(0))<\infty$；$A=A^*\in L(X)$ 为 $\mu$-强单调；$r_0=\lambda^{-1}|F0-b|(1+\sqrt{|A|/\mu})>0$，$t_0=2\lambda/|A|L(r_0)^2$，$0<t<r_0$，则由(5)定义的序列收敛于(1)的唯一解.

**证** 由 $F$ 强单调知(1)有唯一解 $x$. 定义

$$f(y)=(A^{-1}(y-x),y-x),y\in X.$$

令 $\rho=2\lambda t-t^2|A|L(r_0)^2$，则 $\rho>0$. 今归纳地指明

$$|x_k| \leqslant r_0, \rho|x_k - x|^2 \leqslant f(x_k) - f(x_{k+1}). \tag{6}$$

首先，由 $\lambda|x|^2 \leqslant (Fx - F(0), x) \leqslant |F(0) - b||x|$ 得出

$$|x| \leqslant \frac{1}{\lambda}|F(0) - b| \leqslant r_0. \tag{7}$$

利用(5)及 $F$ 强单调不难得出

$$f(x_k) - f(x_{k+1}) \geqslant 2\lambda t|x_k - x|^2 - t^2|A||Fx_k - b|^2. \tag{8}$$

结合(7)(8)得

$$f(x_0) - f(x_1) \geqslant 2\lambda t|x|^2 - t^2|A||F(0) - Fx|^2 \geqslant \rho|x|^2,$$

可见(6)对 $k=0$ 成立. 设(6)对 $0 \leqslant k \leqslant n$ 成立，$n \geqslant 0$. 由 $A$ 自伴及 $\mu|y|^2 \leqslant (Ay, y) \leqslant |A||y|^2$ 推出 $|A|^{-1}|y|^2 \leqslant (A^{-1}y, y) \leqslant \mu^{-1}|y|^2$，于是

$$|A|^{-1}|x_{n+1} - x|^2 \leqslant f(x_{n+1}) \leqslant f(x_n) \leqslant \cdots \leqslant f(x_0)$$

$$= (A^{-1}x, x) \leqslant \mu^{-1}|x|^2 \leqslant \mu^{-1}\lambda^{-2}|F0 - b|^2;$$

$$|x_{n+1}| \leqslant |x_{n+1} - x| + |x| \leqslant \lambda^{-1}|F(0) - b|\left(1 + \sqrt{\frac{|A|}{\mu}}\right) = r_0;$$

$$|Fx_{n+1} - b| = |Fx_{n+1} - Fx| \leqslant L(r_0)|x_{n+1} - x|. \tag{9}$$

结合(8)(9)得出(6)对 $k=n+1$ 成立，于是(6)对 $k \geqslant 0$ 成立. 显然 $f(x_k) \geqslant 0$；(6)推出 $f(x_k)$ 下降，从而收敛，因此(6)推出 $x_k \to x$. □

在 3.9.2 中取 $A=I$ 得到：

**3.9.3 推论** 设 $X$ 为 Hilbert 空间，$F, L(r)$ 如同 3.9.2，$r_0 = 2|F0 - b|/\lambda > 0$，$0 < t < 2\lambda/L(r_0)^2$，$x_0 = 0$，$x_{k+1} = x_k - t(Fx_k - b)$ $(k \geqslant 0)$，则 $\{x_k\}$ 收敛于(1)的唯一解.

**注** 若 $L(r) \leqslant L < \infty$，$0 < t < 2\lambda/L^2$，$X$ 为 Hilbert 空间，$G$ 依(4)，$A=I$，则 $\mathrm{Lip}G \leqslant \sqrt{1 - 2t\lambda + t^2L^2} < 1$，即 $G$ 是压缩映射，此时 3.9.3 中的迭代序列当然收敛.

若 $F: \mathrm{R}^n \to \mathrm{R}^n$ 强单调且为局部 Lip，则 3.9.3 可用. 通常正是利用如 3.9.3 的迭代法解方程(2)，然后用 3.9.1 逼近(1)的解.

对于"梯度算子"3.9.2 有如下 Banach 空间推广：

**3.9.4 定理** 设 $X$ 自反且严格凸；$f: X \to \mathrm{R}$；$F = Df$ 为 $\lambda$-强单

调，$\forall\, r>0$；$L(r)\triangleq \mathrm{Lip}(F|\overline{B}_r(0))<\infty$；$A: X^*\to X$ 为对偶映射，$r_k=|x_k|+|Fx_k-b|$，$2t_k=\min\{1,1/L(r_k)\}$，

$$x_{k+1}=x_k-t_kA(Fx_k-b),\ x_0=0. \tag{10}$$

则 $\{x_k\}$ 收敛于(1)的唯一解 $x$.

**证** 不妨设 $b=0$. 由强单调假定有

$$\lambda|x_k-x|^2\leqslant\langle Fx_k,x_k-x\rangle\leqslant|Fx_k||x_k-x|,$$

故只要证 $Fx_k\to 0$. 直接看出 $|x_k|,|x_{k+1}|\leqslant r_k$. $f$ 为严格凸函数(参考 4.1.5)，用此得出

$$\begin{aligned}f(x_k)-f(x_{k+1})&\geqslant\langle Fx_{k+1},x_k-x_{k+1}\rangle=\lambda_k\langle Fx_{k+1},AFx_k\rangle\\&=\lambda_k\langle Fx_k,AFx_k\rangle+\lambda_k\langle Fx_{k+1}-Fx_k,AFx_k\rangle\\&\geqslant\lambda_k|Fx_k|^2-\lambda_kL(r_k)|x_{k+1}-x_k||Fx_k|\\&=\lambda_k|Fx_k|^2[1-\lambda_kL(r_k)]\geqslant\frac{1}{2}\lambda_k|Fx_k|^2,\end{aligned}$$

这表明 $f(x_k)$ 下降且 $|Fx_k|^2\leqslant 2[f(x_k)-f(x_{k+1})]/\lambda_k$. 由

$$\begin{aligned}f(x_k)-f(x)&=\int_0^1\frac{\mathrm{d}}{\mathrm{d}t}f(x+t(x_k-x))\mathrm{d}t\\&=\int_0^1\langle F(x+t(x_k-x)),x_k-x\rangle\mathrm{d}t\\&\geqslant\int_0^1\lambda t|x_k-x|^2\mathrm{d}t=\frac{\lambda}{2}|x_k-x|^2\end{aligned}$$

推出 $x_k$ 有界，从而 $r_k$ 有界，于是 $\inf\lambda_k>0$，这结合 $f(x_k)$ 的收敛性得出 $Fx_k\to 0$. □

**注** 若 $b=0$，3.9.4 之条件满足，则(1)的解是 $f$ 的唯一极小点(参考 4.1.10). 如 3.9.4 的逼近法称为“梯度法”，而 3.9.2 可看作一种“广义梯度法”.

# 第四章　凸分析与非光滑分析

在现实世界与科学中，非光滑现象经常自然地出现，因而需要有一种处理它们的系统的方法，这导致"非光滑分析"这一数学分支的产生，而"凸分析"则是其中的一个特殊部分．非光滑分析的基本问题是将传统的微分学方法加以推广，使之能在多少类似的形式下用于一定的不可微函数．这一设想已得到辉煌的实现．非光滑分析的结果不仅成功地用于处理非光滑问题，而且也用于那些初看起来没有任何非光滑性的数学问题，且往往导致重要的结论．本章给出非光滑分析的基本理论，它的某些应用包含在相继的章节中．

## §1　凸函数

首先说明有关实泛函的一些术语与记号．鉴于难以避免取值$\infty$的函数，今后说及实泛函时通常指形如$f:X\to\overline{\mathrm{R}}$的函数，其中$\overline{\mathrm{R}}=\mathrm{R}\cup\{\infty\}$．称$D_f\triangleq\{x\in X:f(x)<\infty\}$为$f$的有效域，当$D_f\neq\varnothing$时说$f$是正常的．本书中，任何作为条件给出的实泛函总假定是正常的；至于某一问题讨论过程中出现的实泛函是否正常，则需具体确定．给定$f:D\subset X\to\mathrm{R}$，当必要将其视为$X$上的函数时，常认定$f|D^c=\infty$，从而$D=D_f$．

以下设$f:X\to\overline{\mathrm{R}}$，$D=D_f\neq\varnothing$．称

$$\operatorname{epi}f=\{(x,t)\in X\times\mathrm{R}:f(x)\leqslant t\}\tag{1}$$

为$f$的上图(epigraph)，它是刻划$f$的某些性质的有效工具．若当$x_n\to x$[$x_n\rightharpoonup x$]时$f(x)\leqslant\varliminf\limits_n f(x_n)$，则说$f$在$x$下半连续[弱下半连续]，缩写作lsc[wlsc]；$f$在某集上lsc[wlsc]的意义自明．若$-f$

为 lsc[wlsc]，则说 $f$ 上半连续[弱上半连续]，缩写作 usc[wusc]。不难验证，$f$ 在 $X$ 上 lsc[wlsc]$\Leftrightarrow$epi$f$ 闭[弱闭]。

任给 $t\in J$，本书中常记 $t'=1-t$。若

$$f(tx+t'y)\leqslant tf(x)+t'f(y),0<t<1,x,y\in X, \quad (2)$$

则称 $f$ 为凸函数；若对 $x,y\in D,x\neq y$，(2)是严格不等式，则说 $f$ 严格凸。显然 $f$ 凸$\Leftrightarrow D$ 凸且(2)对 $x,y\in D$ 成立$\Leftrightarrow$epi$f$ 为凸集。若 $f$ 凸，则对任何凸组合 $\sum t_ix_i$ 有 $f(\sum t_ix_i)\leqslant\sum t_if(x_i)$。凸函数是非线性函数中性质特别良好而应用最为广泛的函数类之一，本节概述凸函数的主要性质。

首先指明凸性蕴涵很强的连续性结论。

**4.1.1 定理** 设 $f$ 凸，则以下每个条件推出 $f|D^\circ\in C^{1-0}$：(i)$\exists\ x_0\in D,r,\rho>0$：$f(B_r(x_0))\leqslant\rho$；(ii)$f$ 在某个球 $B_r(x_0)(\subset D)$ 内 lsc；(iii)$f$ 在某点 $x_0\in D$ 连续；(iv)$X=\mathrm{R}^n$。

**证** 设条件(i)满足，不妨设 $x_0=0,f(0)=0$。$\forall\ x\in B_r(0)$，从 $0\leqslant[f(x)+f(-x)]/2\leqslant[f(x)+\rho]/2$ 推出 $f(x)\geqslant-\rho$，因此 $|f(x)|\leqslant\rho$。若 $\delta=r/2$，$x,y\in B_\delta(0)$，$\beta\triangleq|x-y|>0$，则 $z\triangleq x+\beta^{-1}\delta(x-y)\in B_r(0)$，$f(x)=f((\beta+\delta)^{-1}(\beta z+\delta y))\leqslant(\beta+\delta)^{-1}\cdot[\beta f(z)+\delta f(y)]$，于是

$$f(x)-f(y)\leqslant\frac{\beta}{\beta+\delta}[f(z)-f(y)]\leqslant\frac{2\rho}{\delta}|x-y|,$$

故得 $\mathrm{Lip}(f|B_\delta(0))\leqslant2\rho/\delta$。任给 $x\in D^\circ$，取 $k>1$：$kx\in D$，则当 $y$ 邻近 $x$ 时由 $f$ 的凸性有

$$f(y)\leqslant\frac{1}{k}f(kx)+\left(1-\frac{1}{k}\right)f\left(\frac{k(y-x)}{k-1}\right),$$

可见 $f$ 在 $x$ 邻近上有界，从而由已证结论知 $f$ 在 $x$ 邻近亦为 Lip。因此 $f|D^\circ\in C^{1-0}$。

条件(ii)推出 $A_n\triangleq\{x\in\overline{B}_\delta(x_0):f(x)\leqslant n\}$ 闭($0<\delta<r$)。因 $\overline{B}_\delta(x_0)=\bigcup_1^\infty A_n$，故必有某个 $A_n$ 含内点 $x_1$，从而 $f$ 在 $x_1$ 邻近上有界，这表明(ii)$\Rightarrow$(i)。其次显然(iii)$\Rightarrow$(i)。若 $X=\mathrm{R}^n$，$D^\circ\neq\varnothing$，则必

有 $x_i(0\leqslant i\leqslant n)$：$A\triangleq\mathrm{co}\{x_i\}\subset D^\circ$ 且 $A^\circ\neq\varnothing$，显然 $f$ 在 $A$ 内上有界. □

由 4.1.1 之证明顺便得出：

**4.1.2 推论** 若 $f$ 凸，$\forall\ x\in B_{2\delta}(x_0)$：$|f(x)|\leqslant\rho$，则 $\mathrm{Lip}(f|B_\delta(x_0))\leqslant 2\rho/\delta$.

下面讨论凸函数的微分性质.

**4.1.3 命题** 设 $f$ 凸，$x\in D^\circ$. 则单侧方向导数 $D_\pm f(x,z)$ 存在且 $D_-f(x,z)\leqslant D_+f(x,z)\leqslant\Delta f(x,z)\ (\forall\ z\in X)$. 若 $X=\mathrm{R}^n$，$G$-导数 $Df(x)$ 存在，则 $f'(x)=Df(x)$.

**证** 由 $f$ 凸不难推出 $\varphi(t)\triangleq t^{-1}\Delta f(x,tz)$ 在 $t>0$ 时递增，在 $t<0$ 时递降，当 $s<0<t$ 时 $\varphi(s)\leqslant\varphi(t)$，由此直接得出命题的前半部分. 对于后半部分，设 $\{e_i\}$ 是 $\mathrm{R}^n$ 的标准基，$z=\sum z_ie_i\neq 0$，$I=\{i:z_i\neq 0\}$，$t_i=nz_i$. 因 $g\triangleq\Delta f(x,\cdot)-Df(x)$ 凸，$g(tz)=o(t)\ (t\to 0)$，故

$$0\leqslant g(z)=g\left(\sum\frac{t_ie_i}{n}\right)\leqslant\sum\frac{1}{n}g(t_ie_i)=\sum_{i\in I}\frac{z_i}{t_i}g(t_ie_i)$$

$$\leqslant|z|\left\{\sum_{i\in I}\left[\frac{1}{t_i}g(t_ie_i)\right]^2\right\}^{1/2}=o(|z|),$$

由此得出 $f'(x)=Df(x)$. □

**4.1.4 引理** 设 $\varphi:[a,b]\to\mathrm{R}\ (a<b)$. 若 $\varphi$ 可微，则 $\varphi$ 凸[严格凸]$\Leftrightarrow\varphi'$ 增加[严格增加]$\Rightarrow\varphi'$ 在 $(a,b)$ 内连续. 若 $\varphi''$ 存在，则 $\varphi$ 凸 $\Leftrightarrow\varphi''\geqslant 0$；$\varphi''>0\Rightarrow\varphi$ 严格凸.

这是经典分析中的熟知结果. 下面的定理可看作 4.1.4 的一个无限维推广.

**4.1.5 定理** 设 $f$ 在凸开集 $D$ 内 $G$ 可微. (i) $f$ 凸[严格凸]$\Leftrightarrow$ 当 $x,x+z\in D$，$z\neq 0$ 时 $Df(x)z\leqslant\Delta f(x,z)$ $[Df(x)z<\Delta f(x,z)]$ $\Leftrightarrow Df$ 单调[严格单调]. (ii) 若 $f$ 凸，则 $Df$ 半连续，当 $X$ 自反时 $Df$ 次连续，当 $X=\mathrm{R}^n$ 时 $f\in C^1(D)$. (iii) 若 2 阶变分 $D^2f(x,z)\triangleq(\mathrm{d}^2/\mathrm{d}t^2)f(x+tz)|_{t=0}$ 存在，则 $f$ 凸 $\Leftrightarrow D^2f(x,\cdot)\geqslant 0$；$D^2f(x,z)>0\ (x\in D,0\neq z\in X)\Rightarrow f$ 严格凸.

**证** (i)显然 $f$ 凸 $\Rightarrow Df(x)\leqslant\Delta f(x,\cdot)$(4.1.3). 若 $Df(x)\leqslant\Delta f(x,\cdot)(\forall\ x\in D)$,则当 $x,x+z\in D$ 时有

$$\begin{aligned}\langle Df(x+z)-Df(x),z\rangle &= -Df(x+z)(-z)-Df(x)z\\ &\geqslant \Delta f(x,z)-\Delta f(x,z)=0,\end{aligned}$$

可见 $f$ 单调. 若 $Df$ 单调, $x,x+z\in D$, $\varphi(t)=f(x+tz)$,则

$$\varphi'(t)-\varphi'(s)=\langle Df(x+tz)-Df(x+sz),z\rangle\geqslant 0(s<t),$$

由此推出 $\varphi$ 凸(4.1.4),而这推出 $f$ 凸. 关于严格凸的相应结论可类似证明.

(ii)任给 $x\in D$, $y,z\in X$,有

$$\varlimsup_{t\downarrow 0}Df(x+ty)z\leqslant\varlimsup_{t\downarrow 0}\frac{1}{t}\Delta f(x+ty,tz)=Df(x)z.$$

以 $-z$ 换 $z$ 得 $\varliminf_{t\downarrow 0}Df(x+ty)z\geqslant Df(x)z$,故得 $Df(x+ty)\xrightarrow{*}Df(x)$ $(t\downarrow 0)$,即 $Df$ 半连续. 余下结论由 3.1.6, 4.1.3 推出.

(iii)容易从 4.1.4 推出. □

在极值一类的问题中"wlsc"是重要的. 所宜注意者, wlsc 并不能从 lsc(甚至连续)推出,因此其判定颇成问题. 幸而对凸函数有以下结论.

**4.1.6 命题** 设 $f$ 凸. 若 $f$ lsc,或 $f$ 在开集 $D$ 内 $G$-可微,则 $f$ 在 $D$ 内 wlsc.

**证** 若 $f$ lsc,则 $\forall\ r\in\mathrm{R}:\{x:f(x)\leqslant r\}$ 凸闭,从而弱闭[75; 1.8.3],这推出 $f$ 为 wlsc. 若 $f$ $G$-可微, $x_n\rightharpoonup x\in D$,则由 $\langle Df(x),x_n-x\rangle\leqslant f(x_n)-f(x)$(4.1.5)推出

$$f(x)\leqslant\varliminf_{n}[f(x_n)-\langle Df(x),x_n-x\rangle]=\varliminf_{n}f(x_n),$$

可见 $f$ 为 wlsc. □

**4.1.7 命题** 若 $f$ 凸且 lsc,则 $\exists\ \alpha,\beta\geqslant 0:f(x)\geqslant-\alpha|x|-\beta$.

**证** 取 $x_0\in D$, $t_0<f(x_0)$. 因 epi $f$ 闭凸, $(x_0,t_0)\overline{\in}$ epi $f$,故由分离定理有 $(u,r)\in X^*\times\mathrm{R}=(X\times\mathrm{R})^*$,使得

$$u(x_0)+rt_0<u(x)+rt,(x,t)\in\mathrm{epi}\,f. \tag{3}$$

以 $(x,t)=(x_0,f(x_0))$ 代入(3)得 $r>0$；以 $t=f(x)(\forall\ x\in D)$ 代入(3)得 $f(x)>[u(x_0)+rt_0-u(x)]/r$，因此

$$f(x)>-r^{-1}|u||x|-|r^{-1}u(x_0)+t_0|. \qquad \square$$

4.1.7 表明，当 $|x|\to\infty$ 时 lsc 凸函数至多以线性速度下降至 $-\infty$. 这一事实对于某些涉及凸函数的估值是重要的. 亦可注意到，如 $e^x$ 这一简单例子表明的，凸函数可非常快地增长至$\infty$.

凸函数固然有不少良好性质，但对于应用毕竟过于狭窄. 现代研究中愈来愈多地使用某种形式的"广义凸函数"，其主要者定义如下：

**4.1.8 定义** 若$\forall\ x,y\in D$:$f([x,y])\leqslant\max\{f(x),f(y)\}$，则说 $f$ 拟凸；若$\forall\ x,y\in D$: $f(x)\neq f(y)[x\neq y]\Rightarrow f((x,y))<\max\{f(x),f(y)\}$，则说 $f$ 严格拟凸[强拟凸]，注意$(x,y)=\{tx+t'y:0<t<1\}$. 若 $Df(x)$存在，$\forall\ y\in D\backslash\{x\}$: $\langle Df(x),y-x\rangle\geqslant0\Rightarrow f(y)\geqslant f(x)[f(y)>f(x)]$，则说 $f$ 在 $x$ 相对于 $D$ 伪凸[严格伪凸]；$f$ 在 $D$ 上伪凸[严格伪凸]的含义自明.

直接由定义易指明以下结论：$f$ 拟凸$\Leftrightarrow\forall\ r\in\mathrm{R}$: $\{x:f(x)\leqslant r\}$凸($\Leftrightarrow\langle Df(x),y-x\rangle>0$ 蕴涵 $f(y)>f(x)$，若 $f$ $G$-可微)；凸$\Rightarrow$拟凸且严格拟凸；凸且可微$\Rightarrow$伪凸$\Rightarrow$拟凸且严格拟凸；严格凸$\Rightarrow$强拟凸$\Rightarrow$拟凸且严格拟凸；严格凸且可微$\Rightarrow$严格伪凸$\Rightarrow$伪凸.

**4.1.9 例** 1°若 $f:\mathrm{R}\to\overline{\mathrm{R}}$，则 $f$ 单调$\Rightarrow f$ 拟凸；$f$ 严格单调$\Rightarrow f$ 强拟凸；$f$ 可微且 $f'(x)\neq0\Rightarrow f$ 严格伪凸(但未必凸，考虑 $f(x)=\ln x$).

2°设 $x<0\Rightarrow f(x)=0$，$x\geqslant0\Rightarrow f(x)=1$，则 $f$ 拟凸而非严格拟凸. 若 $x\neq0\Rightarrow g(x)=0$，$g(0)=1$，则 $g$ 严格拟凸而非拟凸. 可见"拟凸"与"严格拟凸"互不蕴涵.

各种凸性在与极值有关的问题中都起适当的作用. 以下命题汇集了一些基本的结论.

**4.1.10 命题** 设 $D$ 凸. (i)若 $f$ 严格拟凸，则 $f$ 的局部极小点亦为全局极小点. (ii)若 $f$ 强拟凸，则 $f$ 至多有一个极小点. (iii)若

$X$ 自反,$D$ 闭,$f$ 拟凸、lsc 且弱强制,则 $f$ 有极小点.(iv)若 $\bar{x}$ 是 $f$ 的极小点,则 $D_+f(\bar{x},D-\bar{x})\geqslant 0$;当 $f$ 凸时其逆亦真.(v)若 $f$ 在 $\bar{x}$ 相对于 $D$ 伪凸,则 $\bar{x}$ 是 $f$ 的极小点$\Leftrightarrow\langle Df(\bar{x}),D-\bar{x}\rangle\geqslant 0$($\Leftrightarrow$ $Df(\bar{x})=0$,若 $\bar{x}\in D^\circ$).

**证** 只证(iii)(其余结论的证明是直接的).取 $x_n\in D$,使 $f(x_n)\to\alpha\triangleq\inf\limits_X f(x)$.$x_n$ 必有界,不妨设 $x_n\rightharpoonup\bar{x}$.$D$ 闭凸必弱闭,因此 $\bar{x}\in D$.因 $f$ wlsc(4.1.6),故 $f(\bar{x})\leqslant\alpha$,从而 $\bar{x}$ 是 $f$ 的极小点.

□

本书亦不可避免用到某种凸的向量函数,今陈述其定义以备用.

**4.1.11 定义** 设 $g:D\subset X\to Y$,$Y$ 中由锥 $Y_+$ 导入序$\leqslant$.若$\forall\lambda\in Y_+^*\setminus\{0\}$:$\lambda\circ g$ 凸[严格凸],则说 $g$ 为 * 凸[* 严格凸];"* 伪凸"等的定义仿此;"* 凸"也简称为凸.

若$\forall\lambda\in Y_+^*$:$\lambda\circ g$ 为 lsc,则说 $g$ 为 * lsc.

以下用语是通行的:若 $-f$ 凸,则说 $f$ 凹;严格凹、拟凹等仿此.

## §2 次微分

设 $f:X\to\bar{\mathbb{R}}$,$D=D_f\neq\varnothing$.

**4.2.1 定义** 设 $x\in D$.称 $\partial f(x)\triangleq\{u\in X^*:u\leqslant\Delta f(x,\cdot)\}$为 $f$ 在 $x$ 的次微分,称每个 $u\in\partial f(x)$为 $f$ 在 $x$ 的次梯度;当 $\partial f(x)\neq\varnothing$时说 $f$ 在 $x$ 次可微.

**注** 1°易验知 $u\in\partial f(x)\Leftrightarrow\{(y,t):t-f(x)=u(y-x)\}$是 epi $f$在点$(x,f(x))$的支承超平面,这一事实给出次梯度的几何解释(与导数的切线解释对照).

2°设 $f$ 在凸集 $D$ 上次可微,$z=tx+t'y\in[x,y]\subset D$,取 $u\in\partial f(z)$,则 $f(z)=tf(z)+t'f(z)\leqslant t[f(x)+u(z-x)]+t'[f(y)+$

$u(z-y)]=tf(x)+t'f(y)$，可见 $f$ 凸. 这一事实表明，次微分概念本质上仅适用于凸函数，尽管定义 4.2.1 本身并不涉及函数的凸性.

3°次微分 $\partial f: X\to 2^{X^*}$ 是一多值映射，它正是导数 $f'$ 的某种替代品.

**4.2.2 例** 1°对凸函数 $\varphi: \mathrm{R}\to\mathrm{R}$ 有 $\partial\varphi(x)=[\varphi'_-(x),\varphi'_+(x)]$.

2°设 $f(x)=\frac{1}{2}|x|^2(x\in X)$，今指明 $\partial f=\mathscr{D}$，$\mathscr{D}$ 是 $X$ 的对偶映射. 显然 $\partial f(0)=\mathscr{D}(0)=\{0\}$. 设 $x\neq 0$. 若 $u\in\partial f(x)$，则 $|u||x|=\sup\limits_{|y|=|x|}u(y)=u(x)\leqslant D_+f(x,x)=|x|^2=D_-f(x,x)\leqslant u(x)$，这得出 $u\in\mathscr{D}x$. 反之，若 $u\in\mathscr{D}x$，则 $\forall\ y\in X: u(y-x)\leqslant|u||y|-|x|^2\leqslant f(y)-f(x)$，因此 $u\in\partial f(x)$，故得 $\partial f(x)=\mathscr{D}x$.

3°设 $f(x)=|x|(x\in X)$，则 $x\neq 0$ 时 $\partial f(x)=|x|^{-1}\partial\left(\frac{1}{2}|x|^2\right)=\{x|^{-1}\mathscr{D}x=\{u\in X^*: |u|=1, u(x)=|x|\}$；$u\in\partial f(0)\Leftrightarrow u\leqslant|\cdot|\Leftrightarrow|u|\leqslant 1(u\in X^*)$，因此 $\partial f(0)=\overline{B}_1(0)\subset X^*$.

4°设 $A\subset X, Q\subset X^*$. $A$ 的指示函数 $\delta_A$ 定义为 $\delta_A|A=0, \delta_A|A^c=\infty$. $A$ 与 $Q$ 的支承函数 $s_A$ 与 $s_Q$ 定义为：

$$s_A(u)=\sup\langle u,A\rangle, s_Q(x)=\sup\langle Q,x\rangle(x\in X, u\in X^*). \tag{1}$$

直接看出，$\forall\ x\in A: u\in\partial\delta_A(x)\Leftrightarrow u\in(x-A)^*\Leftrightarrow u(x)=s_A(u)$；总有 $0\in\partial\delta_A(x)$；$x\in A^\circ\Rightarrow\partial\delta_A(x)=\{0\}$；若 $A$ 是锥，则 $\partial\delta_A(x)=(-A)^*\cap\{x\}^\perp$；若 $A$ 是子空间，则 $\partial\delta_A(x)=A^\perp$.

下面给出 $f$ 在一点次可微的条件.

**4.2.3 命题** 设 $f$ 凸，$x\in D$. (i)若 $f$ 在 $x$ 次可微，则 $f$ 在 $x$ lsc. (ii)$u\in\partial f(x)\Leftrightarrow u\leqslant D_+f(x,\cdot)\Leftrightarrow D_-f(x,\cdot)\leqslant u(u\in X^*)$；因此，若 $f$ 在 $x$ 次可微，则 $D_+f(x,\cdot)>-\infty$(这意味着不存在使 $f$ 无限陡地下降的方向). (iii)若$\exists\ r>f(x):(x,r)\in(\mathrm{epi}\ f)^\circ$，则 $f$ 在 $x$ 次可微. (iv)若 $f$ 在 $x$ 连续，则 $f$ 在 $x$ 次可微，且 $\partial f(x)$是弱*紧凸集.

**证** (i)(ii)的证明是直接的(参考 4.1.3).

(iii)由分离定理有$(u,\rho)\in X^*\times\mathrm{R}$,使得$u(x)+\rho f(x)\leqslant u(y)+\rho t$(若$f(y)\leqslant t$),$u(x)+\rho f(x)<u(x)+\rho r$. 由此推出$\rho>0$, $-u/\rho\in\partial f(x)$.

(iv)$f$在$x$连续推出$\{x\}\times(f(x),\infty)\subset(\mathrm{epi}\, f)^\circ$,于是由已证之(iii)知$\partial f(x)\neq\varnothing$. 直接看出$\partial f(x)$为弱$^*$闭凸集. 由$-\Delta f(x,-z)\leqslant u(z)\leqslant\Delta f(x,z)$ $(u\in\partial f(x), z\in X)$及$f$在$x$连续得出$\partial f(x)$等度连续,从而弱$^*$紧. □

**注** 由函数$f=\delta_A$(4.2.2)可见,连续性对于次可微并不是必要的,但 lsc 是必要的.

以下结果指明了次微分与$G$-导数的关系.

**4.2.4 命题** (i)设$f$凸,$x\in D$. 若$Df(x)$存在,则$\partial f(x)=\{Df(x)\}$;若$f$在$x$连续且$\partial f(x)=\{u\}$,则$u=Df(x)$. (ii)若$D$开,$\partial f:D\to X^*$单值半连续,则$\partial f=Df$.

**证** (i)若$Df(x)$存在,则由 4.2.3(ii)得出$\forall\, u\in\partial f(x)$: $Df(x)\leqslant u\leqslant Df(x)$,因此$u=Df(x)$. 其次设$f$在$x$连续且$\partial f(x)=\{u\}$. 取定$0\neq z\in X$, $\forall\ \alpha\in[D_-f(x,z),D_+f(x,z)]$,由$u_\alpha(tz)=\alpha t$定义线性函数$u_\alpha:\mathrm{R}z\to\mathrm{R}$. 由$\alpha$的取法知$u_\alpha(y)\leqslant D_+f(x,y)$ $(y\in\mathrm{R}z)$. 因$D_+f(x,\cdot)$是次线性的,故$u_\alpha$可线性地扩张到$X$上[75;1.7.1],仍记为$u_\alpha$,使得$D_-f(x,\cdot)\leqslant u_\alpha\leqslant D_+f(x,\cdot)\leqslant\Delta f(x,\cdot)$,因此$u_\alpha\in\partial f(x)$(注意$f$在$x$连续并用 4.2.3(ii)). 这推出$u_\alpha=u$,从而$D_\pm f(x,z)=\alpha=u(z)$,于是$Df(x)$存在且$Df(x)=u$.

(ii)取定$x\in D$. $\forall\, z\in X$,当$t>0$充分小时有

$$\langle\partial f(x),tz\rangle\leqslant\Delta f(x,tz)\leqslant\langle\partial f(x+tz),tz\rangle,$$

这推出$D_+f(x,z)=\langle\partial f(x),z\rangle$. 同理$D_-f(x,z)=\langle\partial f(x),z\rangle$,因此$Df(x)=\partial f(x)$. □

§4 中将在更一般的框架内建立“次微分规则”. 此处对凸函数证明以下“和规则”.

**4.2.5 定理**(Moreau-Rockafellar) 设$f,g:X\to\overline{\mathrm{R}}$凸,$f$在某

点 $x\in D_g$ 连续，则 $\partial(f+g)=\partial f+\partial g$.

**证** 取定 $x_0\in D_f\cap D_g$，只需证 $\partial(f+g)(x_0)\subset\partial f(x_0)+\partial g(x_0)$. 不妨设 $x_0=0, f(0)=g(0)=0, 0\in\partial(f+g)(0)$，需证 $0\in\partial f(0)+\partial g(0)$（一般情形可由适当代换得出）. 令 $A=\mathrm{epi}\ f, B=\{(y,t):g(y)\leqslant -t\}$，则 $A,B$ 凸，$A^\circ\neq\varnothing, A^\circ\cap B=\varnothing$，于是由分离定理有 $\omega=(u,\rho)\in X^*\times\mathrm{R}$，使得 $\langle\omega,A\rangle\leqslant\langle\omega,B\rangle$，$\langle\omega,(\hat{x}),f(\hat{x})+1\rangle<\langle\omega,(\hat{x},-g(\hat{x}))\rangle$. 由此易推出 $\rho<0, -u/\rho\in\partial f(0), u/\rho\in\partial g(0)$，因此 $0\in\partial f(0)+\partial g(0)$. □

**注** 4.2.5 的结论显然可推广到任意有限和.

**4.2.6 命题** 设 $\bar{x}\in D$，则 $f(\bar{x})=\min\limits_X f(x)\Leftrightarrow 0\in\partial f(\bar{x})\Leftrightarrow 0\in\partial f(\bar{x})+\partial\delta_D(\bar{x})\Leftrightarrow\partial f(\bar{x})\cap(D-\bar{x})^*\neq\varnothing$；若 $D$ 是子空间，则 $f(\bar{x})=\min\limits_X f(x)\Leftrightarrow\partial f(\bar{x})\cap D^\perp\neq\varnothing$.

证明是直接的（参考 4.2.2 之 4°）. 4.2.6 推广了熟知极值条件：若 $\bar{x}$ 是 $f$ 的极小点且 $f'(\bar{x})$ 存在，则 $f'(\bar{x})=0$. 所宜注意者，对于 $\bar{x}$ 为 $f$ 的全局极小点，$0\in\partial f(\bar{x})$ 不仅必要而且是充分的.

**4.2.7 定理**（Rockafellar, 1970） 设 $f$ 凸且 lsc，则 $\partial f: X\to 2^{X^*}$ 是极大单调算子（3.1.7）.

**证** 可直接验知 $\partial f$ 单调. 下面仅就 $X$ 自反的情况证明 $\partial f$ 极大单调. 可设 $X$ 与 $X^*$ 严格凸（3.2.4），只需证 $R(\partial f+\mathscr{D})=X^*$（3.3.5）. $\forall\, u\in X^*$，令 $g=f+\frac{1}{2}|\cdot|^2-u$，则 $g$ 凸、lsc，且 $\lim\limits_{|x|\to\infty}g(x)=\infty$（用 4.1.7），于是有极小点 $\bar{x}\in D$（4.1.10），因此 $0\in\partial g(\bar{x})=\partial f(\bar{x})+\mathscr{D}\bar{x}-u$（4.2.6，4.2.5，4.2.2），这表明 $u\in(\partial f+\mathscr{D})(\bar{x})$. □

## §3 Clarke 次微分

本节设 $\Omega\subset X$ 开，$f:\Omega\to\mathrm{R}$，$\mathrm{Lip}f\leqslant K<\infty$.

**4.3.1 定义** 设 $x\in\Omega$. 称

$$f^\circ(x,z)\triangleq\overline{\lim_{y\to x,t\downarrow 0}}\frac{1}{t}\Delta f(y,tz)\quad(z\in X)\tag{1}$$

为 $f$ 在 $x$ 的 Clarke 广义方向导数;称

$$\partial f(x)\triangleq\partial_z f^\circ(x,0)=\{u\in X^*:u\leqslant f^\circ(x,\cdot)\}\tag{2}$$

为 $f$ 在 $x$ 的 Clarke 次微分(或广义梯度). 若 $f^\circ(x,\cdot)=D_+f(x,\cdot)$,则说 $f$ 在 $x$ 是正则的.

**注** 1° $f^\circ(x,z)$ 必存在且 $|f^\circ(x,z)|\leqslant K|z|$. 用(1)易验知 $f^\circ(x,\cdot)$ 是次线性的,因而为连续凸函数(4.1.1),于是由(2)与 4.2.3 知 $\partial f(x)$ 为非空弱*紧凸集.

2° 若 $f$ 凸,$x\in\Omega$,则 $f$ 必在 $x$ 正则:

$$D_+f(x,z)\leqslant f^\circ(x,z)\leqslant\lim_{t\downarrow 0}\lim_{y\to x}\frac{1}{t}\Delta f(y,tz)=D_+f(x,z).$$

于是对照(2)与 4.2.3(ii)知 $\partial f(x)$ 重合于 §2 中的次微分.

3° 若 $x$ 是 $f$ 的局部极小(或极大)点,则必 $f^\circ(x,\cdot)\geqslant 0$,从而 $0\in\partial f(x)$. 用于表达极值条件,乃是引进广义梯度的基本考虑之一.

以下定理汇集了映射 $f^\circ:\Omega\times X\to\mathrm{R}$ 与 $\partial f:\Omega\to 2^{X^*}$ 的主要性质,这些性质今后将反复引用.

**4.3.2 定理** $f^\circ$ 与 $\partial f$ 有以下性质:(i) $f^\circ$ 为 usc; $f^\circ(x,-z)=(-f)^\circ(x,z)$; $\mathrm{Lip}f^\circ(x,\cdot)\leqslant K$; $\mathrm{R}(\partial f)\subset\overline{B}_K(0)\subset X^*$. (ii) $\partial f$ 是 sw*-闭的(3.1.1). (iii) 若 $\varnothing\neq Q\subset X^*$ 弱*紧凸,则 $\partial f(x)\subset Q\Leftrightarrow f^\circ(x,\cdot)\leqslant s_Q$, $\partial f(x)=Q\Leftrightarrow f^\circ(x,\cdot)=s_Q$, $s_Q$ 依 §2(1). (iv) 若 $X$ 自反,则 $\partial f$ 为 sw*-usc;若 $X=\mathrm{R}^n$,则 $\partial f$ 为 usc(3.1.1).

**证** (i)之证是直接的;用 $f^\circ(\cdot,z)$usc 推出(ii).

(iii) 令 $P=\partial f(x)$,则 $s_P\leqslant f^\circ(x,\cdot)$. $\forall z_0\in X$,由 $u(tz_0)=tf^\circ(x,z_0)$ 定义 $u:\mathrm{R}z_0\to\mathrm{R}$. 如 4.2.4 之证,可设 $u\in X^*$ 且 $u\leqslant f^\circ(x,\cdot)$,于是 $u\in P$, $u(z_0)=f^\circ(x,z_0)$. 故得 $s_P=f^\circ(x,\cdot)$. 若 $P\subset Q$,则 $f^\circ(x,\cdot)=s_P\leqslant s_Q$. 若 $P\not\subset Q$,则 $\exists\ u\in P\backslash Q$,由分离定理有 $z\in$

$X:\langle Q,z\rangle<u(z)$，从而 $f^\circ(x,z)\leqslant s_Q(z)$. 类似地可证 $P=Q\Leftrightarrow s_P=s_Q$.

(iv)若 $\partial f$ 在某点 $x\in\Omega$ 非 sw*-usc，则∃ $\varepsilon>0, z\in X, x_n\to x$, $u_n\in\partial f(x_n)$，使得∀ $u\in\partial f(x)$: $|\langle u_n-u,z\rangle|\geqslant\varepsilon$. 因 $|u_n|\leqslant K$，不妨设 $u_n\xrightarrow{*}u$. 由(ii)有 $u\in\partial f(x)$，这与 $|\langle u_n-u,z\rangle|\geqslant\varepsilon$ 矛盾. □

4.3.2(iii)蕴涵等式 $f^\circ(x,z)=\max\langle\partial f(x),z\rangle$. 此公式与(2)一起表明 $f^\circ$ 与 $\partial f$ 相互唯一确定.

若 $F:\Omega\to Y$ 在 $x$ 点 $G$-可微，对任何紧集 $C\subset X$，关于 $z\in C$ 一致地成立

$$DF(x)z=\lim_{y\to x,t\downarrow 0}\frac{1}{t}\Delta F(y,tz),\tag{3}$$

则说 $F$ 在 $x$ 严格可微，称 $DF(x)$ 为严格导数. 记作 $D_sf(x)$.

**4.3.3 命题** 设(3)对任何 $z\in X$ 成立，则 $F$ 在 $x$ 严格可微⇔$F$ 在 $x$ 邻近为 Lip.

**证** 若 $F$ 在 $x$ 邻近不为 Lip，则有 $x_n, y_n\to x$, $|x_n-y_n|<1/n$, $|Fx_n-Fy_n|>n|x_n-y_n|$. 令 $t_n=\sqrt{n}\,|x_n-y_n|$, $z_n=(y_n-x_n)/t_n$，则 $t_n\to 0, z_n\to 0$, $|\Delta F(x_n,t_nz_n)/t_n|>\sqrt{n}$，这表明(3)不对 $z=z_n$ 一致成立. 反之，若 $DF(x)$ 非严格导数，则有 $z_n\to z, x_n\to x, t_n\to 0$，使 $|t_n^{-1}\Delta F(x_n,t_nz_n)-DF(x)z_n|\geqslant\varepsilon>0$，于是 $F$ 在 $x$ 邻近必不为 Lip，否则将有

$$\begin{aligned}\varepsilon\leqslant&|t_n^{-1}||F(x_n+t_nz_n)-F(x_n+t_nz)|\\&+|t_n^{-1}\Delta F(x_n,t_nz)-DF(x)z|\\&+|DF(x)z-DF(x)z_n|\to 0(n\to\infty)!\end{aligned}$$ □

利用严格导数概念，可得 4.2.4 的如下类似.

**4.3.4 命题** 若 $Df(x)$ 存在，则 $Df(x)\in\partial f(x)$; $\partial f(x)=\{u\}\Leftrightarrow u=D_sf(x)\Leftrightarrow u=Df(x)$ 且 $f$ 在 $x$ 正则；若 $f\in C^1$，则 $\partial f(x)=f'(x)=D_sf(x)$；若 $X=\mathrm{R}^n$, $\partial f$ 单值，则 $f\in C^1$.

**证** 若 $\partial f(x)=\{u\}$，则 $f^\circ(x,z)=u(z)$(4.3.2(iii)). 而

$$\lim_{y\to x,t\downarrow 0}\frac{1}{t}\Delta f(y,tz)=-f^\circ(x,-z)=u(z),$$

故由 4.3.3 有 $u=D_sf(x)$. 其余结论是明显的. □

**注** 若 $f$ 可微而非严格可微,则不必 $\partial f(x)=f'(x)$. 例如对 $f(x)=x^2\sin x^{-1}$, $f(0)=0$,有 $\partial f(0)=[-1,1]$,而 $f'(0)=0$!

**4.3.5 中值定理**(Leboug,1975) 若 $[x,y]\subset\Omega$,则

$$f(y)-f(x)\in\langle\partial f([x,y]),y-x\rangle. \tag{4}$$

**证** 令 $z_t=t'x+ty$, $g(t)=f(z_t)$, $\varphi(t)=g(t)+t[f(x)-f(y)]$,则 $\varphi(0)=f(x)=\varphi(1)$,于是 $\varphi$ 有极值点 $t\in(0,1)$,从而 $0\in\partial\varphi(t)$. 直接验知 $\partial\varphi(t)=\partial g(t)+f(x)-f(y)$,于是只需证 $\partial g(t)\subset\langle\partial f(z_t),y-x\rangle\triangleq Q$, $Q\subset R$ 必为紧区间. 由 4.3.2,这归于证 $\forall\, r\in\mathrm{R}$: $g^\circ(t,r)\leqslant\max rQ$,这由明显的不等式 $g^\circ(t,r)\leqslant f^\circ(z_t,r(y-x))$ 推出. □

**4.3.6 推论** 若 $\Omega$ 凸,则 $f$ 凸 $\Leftrightarrow\partial f$ 单调(参看 4.1.5,4.2.7).

**证** 设 $\partial f$ 单调, $x_1,x_2\in\Omega$, $z=\sum t_ix_i\in[x_1,x_2]$. 取 $y_i\in[x_i,z]$, $u_i\in\partial f(y_i)$,使 $f(x_i)-f(z)=\langle u_i,x_i-z\rangle(i=1,2)$,则

$$\sum t_if(x_i)-f(z)=\rho\langle u_1-u_2,y_1-y_2\rangle\geqslant 0,$$

其中 $\rho=t_1|x_1-z|/|y_1-y_2|\geqslant 0$. 由此可见 $f$ 凸. □

下面考虑次微分对于极值问题的应用. 前已指明,若 $x\in\Omega$ 是 $f$ 的极值点,则 $0\in\partial f(x)$. 若 $x$ 仅是 $f|D$ 的极小点, $D\subset\Omega$ 不必为开集,则条件"$0\in\partial f(x)$"需作适当修改. 可考虑求助于指示函数 $\delta_D$(参考 4.2.6 及下面的 4.3.8);更精细的方法是使用"距离函数"$d_D(x)\triangleq d(x,D)$. 与 $\delta_D$ 比较, $d_D$ 包含更多的信息且有较好的性质,因而是非光滑分析中一个颇为有用的工具.

容易直接验证: $d_D(x)=0\Leftrightarrow x\in\overline{D}$; $D$ 凸 $\Rightarrow d_D$ 凸 $\Rightarrow\overline{D}$ 凸; $\mathrm{Lip}\,d_D\leqslant 1$, $R(\partial d_D)\subset\overline{B}_1(0)\subset X^*$; $x\in D\Rightarrow 0\in\partial d_D(x)$. 这些性质下面将不加指明地反复使用.

**4.3.7 定理** 设 $D\subset\Omega$, $\bar{x}$ 是 $f|D$ 的极小点, $\varphi_\alpha=f+\alpha d_D$. (i)

若 $\alpha \geqslant K$，则 $\bar{x}$ 是 $\varphi_\alpha$ 的极小点，

$$0 \in \partial f(\bar{x}) + \alpha \partial d_D(\bar{x}). \tag{5}$$

(ii)若 $\alpha > K$，$D$ 闭，则 $f|D$ 与 $\varphi_\alpha$ 有同样的极小点.

**证** (i)若 $\bar{x}$ 非 $\varphi_\alpha$ 之极小点，则∃ $x \in \Omega, \varepsilon > 0$：$\varphi_\alpha(x) < f(\bar{x}) - \alpha\varepsilon$. 取 $y \in D$，使 $|x-y| < d_D(x) + \varepsilon$，则

$$f(y) \leqslant f(x) + K|y - x| \leqslant f(x) + \alpha d_D(x) + \alpha\varepsilon < f(\bar{x}),$$

得出矛盾. 因此 $0 \in \partial\varphi_\alpha(\bar{x})$，而这得出(5).

(ii)任取 $\varphi_\alpha$ 的极小点 $x$，今证 $x \in D$(从而 $x$ 是 $f|D$ 的极小点). 令 $\beta = (\alpha + K)/2$，则 $K < \beta < \alpha$，于是 $\varphi_\alpha(x) = \varphi_\alpha(\bar{x}) = \varphi_\beta(\bar{x}) \leqslant \varphi_\beta(x)$，$\alpha d_D(x) \leqslant \beta d_D(x)$，这推出 $d_D(x) = 0, x \in D$. □

(5)可与 4.2.6 中的条件

$$0 \in \partial f(\bar{x}) + \partial\delta_D(\bar{x}) \tag{6}$$

对照. 实际上，当 $\partial f(\bar{x})$ 是 Clarke 次微分时(6)亦可用作 $f|D$ 的极值条件. 为表述下之结果，引入概念：若 $x \in \Omega, \forall\ y \in D$：$f^\circ(x, y-x) \geqslant 0 \Rightarrow f(y) \geqslant f(x)$，则说 $f$ 在 $x$ 相对于 $D$ 为广义伪凸(参照 4.1.8).

**4.3.8 定理** 设 $\bar{x} \in D \subset \Omega$，$D$ 凸. 若 $\bar{x}$ 是 $f|D$ 的极小点，则(6)成立；当 $f$ 在 $\bar{x}$ 相对于 $D$ 为广义伪凸时其逆亦真.

**证** 首先设 $\bar{x}$ 是 $f|D$ 的极小点. 若(6)不真，则 $\partial f(\bar{x}) \cap (D-\bar{x})^* = \varnothing$(参考 4.2.2). 因 $\partial f(\bar{x})$ 弱*紧凸而 $(D-\bar{x})^*$ 弱*闭凸，故由分离定理有 $z \in X$：$\langle \partial f(\bar{x}), z\rangle < \langle (D-\bar{x})^*, z\rangle$. 这推出 $f^\circ(\bar{x}, z) = \max\langle \partial f(\bar{x}), z\rangle < 0$，而 $\langle (D-\bar{x})^*, z\rangle \geqslant 0$，后者推出 $z \in \overline{R_+(D-\bar{x})}$(用 2.2.3). 取 $\alpha > 0, x \in D$，使 $\alpha(x-\bar{x})$ 充分邻近 $z$，则 $f^\circ(\bar{x}, \alpha(x-\bar{x})) = \alpha f^\circ(\bar{x}, x-\bar{x}) < 0$. 另一方面(注意 $D$ 凸！)

$$f^\circ(\bar{x}, x - \bar{x}) \geqslant \varlimsup_{t \downarrow 0} \frac{1}{t} \Delta f(\bar{x}, t(x - \bar{x})) \geqslant 0,$$

得出矛盾.

反之，若(6)成立，则∃ $u \in \partial f(\bar{x}) \cap (D-\bar{x})^*$，于是 $\forall\ x \in D$：$f^\circ(\bar{x}, x-\bar{x}) \geqslant \langle u, x-\bar{x}\rangle \geqslant 0$. 因此，当 $f$ 在 $\bar{x}$ 相对于 $D$ 广义伪凸

时 $f(D)\geqslant f(\bar{x})$. □

## §4 次微分规则

如同经典微分学中一样，基本的次微分规则是"链规则"，其它许多规则由其导出.

**4.4.1 定理** 设 $f=\varphi\circ g, \varphi: \mathrm{R}^m\to\mathrm{R}, g=(g_i): X\to\mathrm{R}^m, g,\varphi$ 分别在 $x\in X$ 与 $y=g(x)$ 邻近为 Lip，则

$$\partial f(x)\subset\overline{\mathrm{co}}\partial\varphi(y)\circ\partial g(x), \tag{1}$$

其中 $\overline{\mathrm{co}}$ 记弱 $^*$ 闭凸包(下同)，$\partial g(x)\triangleq\Pi\partial g_i(x)\subset L(X,\mathrm{R}^m)$. 以下每个条件推出(1)为等式：(i) $g_i$ 在 $x$ 及 $\varphi$ 在 $y$ 正则，$\partial\varphi(y)\subset\mathrm{R}^m_+$；(ii) $m=1$，$\varphi$ 在 $y$ 严格可微；(iii) $g$ 在 $x$ 严格可微，$\varphi$ 在 $y$ 正则. 在条件(i)或(iii)之下 $f$ 在 $x$ 正则；在条件(ii)或(iii)之下可去掉记号 $\overline{\mathrm{co}}$.

**证** 令 $Q=\partial\varphi(y)\circ\partial g(x)$，则 $Q$(从而 $\overline{\mathrm{co}}Q$)弱 $^*$ 紧，只要证 $f^\circ(x,\cdot)\leqslant s_Q$(4.3.2). 取定 $z\in X$. $\forall\,\varepsilon>0$，令

$$q_\varepsilon=\sup\left\{\sum\alpha_i u_i(z);\alpha=(\alpha_i)\in\partial\varphi(B_\varepsilon(y)),u_i\in\partial g_i(B_\varepsilon(x))\right\}, \tag{2}$$

$q=s_Q(z)$，只要证 $f^\circ(x,z)<q_\varepsilon+\varepsilon$，$q_\varepsilon\to q(\varepsilon\downarrow 0)$. 取 $x'\in B_\varepsilon(x)$，$t>0$，使 $f^\circ(x,z)<t^{-1}\Delta f(x',tz)+\varepsilon$，且 $x'+tz\in B_\varepsilon(x)$，$|g(x')-y|<\varepsilon$，$|g(x'+tz)-y|<\varepsilon$. 由 4.3.5，有

$$\begin{aligned}\Delta f(x',tz)&\in\langle\partial\varphi([g(x'),g(x'+tz)]),\Delta g(x',tz)\rangle\\&\subset\langle\partial\varphi(B_\varepsilon(y)),\Delta g(x',tz)\rangle;\end{aligned}$$

$$\Delta g_i(x',tz)\in\langle\partial g_i(B_\varepsilon(x)),tz\rangle.$$

对照(2)看出 $\Delta f(x',tz)\leqslant tq_\varepsilon$，从而 $f^\circ(x,z)<q_\varepsilon+\varepsilon$. 不妨设 $\mathrm{Lip}\varphi$，$\mathrm{Lip}g\leqslant K<\infty$. $\forall\,\delta>0$，取 $\varepsilon>0$ 充分小，使 $\partial\varphi(B_\varepsilon(y))\subset\partial\varphi(y)+B_\delta(0)$，且 $g^\circ_i(B_\varepsilon(x),\pm z)\leqslant g^\circ_i(x,\pm z)+K^{-1}\delta$(4.3.2)，则

$$\begin{aligned}q_\varepsilon&\leqslant\sup\left\{\sum\sup\{\alpha_i u_i(z);u_i\in\partial g_i(B_\varepsilon(x))\};\alpha\in\partial\varphi(B_\varepsilon(y))\right\}\\&\leqslant\sup\left\{\sum[g^\circ_i(x,\alpha_i z)+\delta];\alpha\in\partial\varphi(B_\varepsilon(y))\right\}\end{aligned}$$

$$\leqslant \sup\Big\{\sum \max\langle \partial g_i(x), \alpha_i z\rangle : \alpha \in \partial\varphi(y) + B_\delta(0)\Big\} + m\delta$$

$$\leqslant q + m\delta(K|z| + 1),$$

这与 $q \leqslant q_\varepsilon$ 一起得出 $q_\varepsilon \to q(\varepsilon \downarrow 0)$. 于是(1)式成立.

若条件(i)满足,则依上段记号有

$$\begin{aligned} q &= \max\Big\{\sum \max\langle \partial g_i(x), \alpha_i z\rangle : \alpha \in \partial\varphi(y)\Big\} \\ &= \max\Big\{\sum \alpha_i g^\circ_i(x,z) : \alpha \in \partial\varphi(y)\Big\} \\ &= \max\langle \partial\varphi(y), \omega\rangle \quad (\omega = (D_+ g_i(x,z))) \\ &= \varphi^\circ(y,\omega) = D_+ \varphi(y,\omega) = D_+ f(x,z), \end{aligned}$$

这表明 $q = f^\circ(x,z) = D_+ f(x,z)$, 于是(1)是等式且 $f$ 在 $x$ 正则. 当条件(iii)满足时可类似处理. 若条件(ii)满足,可设 $\alpha \triangleq D_s\varphi(y) \geqslant 0$,则

$$q = \alpha g^\circ(x,z) = \overline{\lim_{\xi\to x, t\downarrow 0}} \frac{\alpha}{t}\Delta g(\xi, tz) = f^\circ(x,z),$$

这同样得出(1)为等式. □

**4.4.2 定理**(Aubin-Clarke, 1979) 设 $f = \varphi \circ g$, $\varphi: Y \to \mathrm{R}$, $g: X \to Y$, $g$ 在 $x \in X$ 严格可微, $\varphi$ 在 $y = g(x)$ 邻近为 Lip, 则

$$\partial f(x) \subset \partial\varphi(y) \circ Dg(x). \tag{3}$$

以下每个条件推出(3)为等式:(i) $\varphi$ 或 $-\varphi$ 在 $y$ 正则;(ii) $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$: $B_\delta(y) \subset \overline{gB_\varepsilon(x)}$;(iii) $R(Dg(x)) = Y$.

**证** 令 $A = Dg(x)$, $Q = A^* \partial\varphi(y)$, 则(3)相当于 $f^\circ(x, \cdot) \leqslant s_Q$(4.3.2),后者由以下推演得出:

$$\begin{aligned} f^\circ(x,z) &= \overline{\lim_{\xi\to x, t\downarrow 0}} \frac{1}{t}[\varphi(g(\xi + tz)) - \varphi(g(\xi))] \\ &\leqslant \varphi^\circ(y, Az) + \overline{\lim_{\xi\to x, t\downarrow 0}} \frac{1}{t}[\varphi(g(\xi + tz)) - \varphi(g(\xi) + tAz) \\ &\leqslant \varphi^\circ(y, Az) + \overline{\lim_{\xi\to x, t\downarrow 0}} \mathrm{const}\left|\frac{1}{t}\Delta g(\xi + tz) - Az\right| \\ &= \varphi^\circ(y, Az) = \max\langle \partial\varphi(y), Az\rangle = s_Q(z), z \in X. \end{aligned}$$

若 $\varphi$ 在 $y$ 正则($-\varphi$ 正则时类此),则

$$\varphi^{\circ}(y,Az) = D_+\,\varphi(y,Az) = D_+\,f(x,z),$$

这得出 $s_Q(z)=D_+f(x,z)=f^{\circ}(x,z)(\forall\ z\in X)$，因此(3)为等式且 $f$ 在 $x$ 正则. 若条件(ii)满足，则

$$\varphi^{\circ}(y,Az) = \overline{\lim_{\eta\to y,t\downarrow 0}}\ \frac{1}{t}\Delta\varphi(\eta,tAz)$$

$$\leqslant \overline{\lim_{\xi\to x,t\downarrow 0}}\ \frac{1}{t}\Delta\varphi(g(\xi),tAz) = f^{\circ}(x,z),$$

这同样得出 $s_Q(z)=f^{\circ}(x,z)$. 当 $R(A)=Y$ 时(ii)必满足. □

4.4.1 与 4.4.2 有一系列推论，其中首先包括熟知微分规则的以下推广：

**4.4.3 推论** 设 $g,g_i:X\to\mathrm{R}$ 在点 $x$ 邻近为 Lip，则

$$\partial(\alpha g)(x) = \alpha\partial g(x),\alpha\in\mathrm{R}; \tag{4}$$

$$\partial\Big(\sum_{i=1}^{m}g_i\Big)(x) \subset \sum_{i=1}^{m}\partial g_i(x); \tag{5}$$

$$\partial(g_1g_2)(x) \subset g_2(x)\partial g_1(x) + g_1(x)\partial g_2(x); \tag{6}$$

$$\partial(g_1/g_2)(x) \subset [g_2(x)\partial g_1(x) - g_1(x)\partial g_2(x)]/g_2^2(x). \tag{7}$$

当 $g_i(1\leqslant i\leqslant m)$ 在 $x$ 正则时(5)是等式；当 $g_i$ 在 $x$ 正则且 $g_i(x)\geqslant 0$ $(i=1,2)$ 时(6)是等式；(7)要求 $g_2(x)\neq 0$；当 $g_1$ 与 $-g_2$ 在 $x$ 正则且 $g_1(x)\geqslant 0,g_2(x)>0$ 时(7)是等式.

以上结论可直接从 4.4.1 推出. 例如，令 $g=(g_i),\varphi(y_1,\cdots,y_m)=\sum y_i$，对 $\sum g_i=\varphi\circ g$ 应用 4.4.1 得出(5).

**4.4.4 推论** 设 $A\in L(X,Y),x\in X,\varphi:Y\to\mathrm{R}$ 在 $Ax$ 邻近为 Lip 且在 $Ax$ 正则，则 $\partial(\varphi\circ A)(x)=A^*\partial\varphi(Ax)$.

此由 4.4.2 推出. 注意连续凸(或凹)的 $\varphi$ 满足 4.4.4 之条件.

**4.4.5 推论** 设 $X$ 连续并稠密地嵌入 $Y,x\in X,\varphi:Y\to\mathrm{R}$ 在 $x$ 邻近为 Lip，则 $\partial(\varphi|X)(x)=\partial\varphi(x)$，这意味着每个 $u\in\partial(\varphi|X)(x)$ 可唯一地扩张为 $\partial\varphi(x)$ 中一元.

这由应用 4.4.2 于 $\varphi|X=\varphi\circ g$ 得出，此处 $g:X\to Y$ 是嵌入映射.

下面设 $\varphi: X=X_1\times X_2\to\mathrm{R}$，分别以 $\partial_i\varphi$ 与 $\varphi_i^\circ$ 记 $\varphi(x_1,x_2)$ 对 $x_i$ 的广义梯度与广义方向导数，今考虑"偏次微分"$\partial_i\varphi$ 与 $\partial\varphi$ 之关系.

**4.4.6 定理** 设 $\varphi$ 在 $x=(x_1,x_2)$ 邻近为 Lip. (i) $\partial_i\varphi(x)\subset\pi_i\partial\varphi(x)$，$\pi_i: X^*\to X_i^*$ 是投影，$i=1,2$. (ii)若 $\varphi$ 在 $x$ 正则，则 $\partial_i\varphi(x)=\pi_i\partial\varphi(x)$，从而 $\partial\varphi(x)\subset\partial_1\varphi(x)\times\partial_2\varphi(x)$.

**证** (i)对 $\varphi(\cdot,x_2)=\varphi\circ g\,(g=(\cdot,x_2))$ 应用 4.4.2 得 $\partial_1\varphi(x)\subset\pi_1\partial\varphi(x)$；同理 $\partial_2\varphi(x)\subset\pi_2\partial\varphi(x)$.

(ii)任给 $u=(u_1,u_2)\in\partial\varphi(x)$，$u_i\in X_i^*\ (i=1,2)$，今证 $u_1\in\partial_1\varphi(x)$（同理将有 $u_2\in\partial_2\varphi(x)$）. $\forall\ z\in X_1$，有

$u_1(z)=u(z,0)\leqslant\varphi^\circ(x,(z,0))=D_+\varphi(x,(z,0))\leqslant\varphi_1^\circ(x,z)$，

这正表明 $u_1\in\partial_1\varphi(x)$. □

上面建立的"次微分规则"无疑有助于次微分的计算. 然而，在一般情况下，次微分计算远不是容易的. 关键在于，$\partial f(x)$ 的定义（§3(2)）不是构造性的；$f^\circ(x,\cdot)$ 的定义（§3(1)）虽然具有构造性，但§3(1)并不是一个易操作的算式. 下面就 $X=\mathrm{R}^n$ 的情况给出 $\partial f(x)$ 的一个构造性表示.

**4.4.7 定理** 设 $f:\Omega\subset\mathrm{R}^n\to\mathrm{R}$ 为 $C^{1-0}$ 函数，$S\subset\mathrm{R}^n$ 是(Lebesgue)零测集，$N$ 记 $f$ 的不可微点集. 则

$$\partial f(x)=\mathrm{co}\{\lim_k f'(x_k): x_k\to x, x_k\,\overline{\in}\,S\cup N\}.\ (x\in\Omega)\quad(8)$$

约定"$f'(x_k)$收敛"这一限制已隐含于表达式中.

**证** 由一个著名定理（参考[29]），$f$ 几乎处处可微，因此 $\mathrm{mes}N=0$. 以 $Q$ 记(8)之右端，则易见 $Q$ 是非空有界凸闭集，且 $Q\subset\partial f(x)$（4.3.3，4.3.2(ii)），于是只要证 $f^\circ(x,\cdot)\leqslant s_Q$（4.3.2(iii)）. 为此只需证：若 $\varepsilon>0$，$0\neq z\in\mathrm{R}^n$，则

$$f^\circ(x,z)-\varepsilon\leqslant\varliminf_{y\overline{\in}S\cup N,\,y\to x}\langle f'(y),z\rangle\triangleq\beta.\qquad(9)$$

可设 $|z|=1$. 取 $\delta>0$，使得 $\forall\ y\in B_{2\delta}(x)\backslash(S\cup N)$：$\langle f'(y),z\rangle<\beta+\varepsilon$. 由 Fubini 定理，对几乎所有 $y\in B_\delta(x)$，$[y,y+\delta z]\cap(S\cup N)$ 是 1 维零测集；对这样的 $y$ 与任何 $t\in(0,\delta)$ 有

$$\frac{1}{t}\Delta f(y,tz) = \frac{1}{t}\int_0^t \langle f'(y+sz),z\rangle \mathrm{d}s \leqslant \beta + \varepsilon.$$

由连续性，不等式 $t^{-1}\Delta f(y,tz)\leqslant\beta+\varepsilon$ 可用于任何 $y\in B_\delta(x)$ 与 $t\in(0,\delta)$，而这意味着(9)成立. □

若令 $B_k=B(x,1/k)\setminus(s\cup N)$，则(8)式可表成：

$$\partial f(x) = \mathrm{co}\bigcap_{k=1}^{\infty}\overline{f'(B_k)}. \tag{10}$$

当 $f'(B_k)$ 不太复杂时应用(10)是方便的.

**4.4.8 例** 1°设 $f(x,y)=\max\{\min\{x,-y\},y-x\}$，$(x,y)\in \mathrm{R}^2$，则

$$f(x,y)=\begin{cases} x, & y\leqslant 2x \text{ 且 } y\leqslant -x; \\ -y, & -2x\leqslant 2y\leqslant x; \\ y-x, & x\leqslant 2y \text{ 或 } 2x\leqslant y. \end{cases}$$

直接看出 $f'(x,y)$ 仅取值 $(1,0)$，$(0,-1)$，$(-1,1)$，以此三点为顶点的实心三角形即 $\partial f(0,0)$. 其次易知 $\partial_x f(0,0)=[-1,0]$，$\partial_y f(0,0)=[0,1]$，可见 $\partial f(0,0)\not\subset\partial_x f(0,0)\times\partial_y f(0,0)$，从而 $f$ 在 $(0,0)$ 不是正则的(4.4.6).

2°设 $f(x)=x^2\sin x^{-1}$，$f(0)=0$，则 $f$ 处处可微，且 $f'(x)=2x\sin x^{-1}-\cos x^{-1}$ $(x\neq0)$，$f'(0)=0$. 这结合(10)得出 $\partial f(0)=[-1,1]$. 这表明 $f$ 在 $x=0$ 可微但非严格可微(4.3.4).

今将 4.4.7 用于求 $\partial d_D(x)$. 首先引入概念：若 $D\subset\mathrm{R}^n$，$0\neq z\in\mathrm{R}^n$，$z$ 可表为 $x'-x$，$x$ 是 $\overline{D}$ 中唯一最接近 $x'$ 的点，则说 $z$ 在 $x$ 垂直于 $D$，记作 $z\perp_x D$.

**4.4.9 引理** 设 $D\subset\mathrm{R}^n$，$x\in\mathrm{R}^n$，$d'_D(x)\neq0$，则 $|d'_D(x)|=1$，且$\exists$ $y\in\overline{D}$：$d'_D(x)=(x-y)/|x-y|$，$|x-y|d'_D(x)\perp_y D$.

**证** 取 $y\in\overline{D}$，使 $d_D(x)=|x-y|$. 则

$$\begin{aligned}\langle d'_D(x),x-y\rangle &= \lim_{t\downarrow 0}\frac{1}{t}[d_D(x)-d_D(x+t(y-x))] \\ &= |x-y|. \end{aligned} \tag{11}$$

必定 $x\neq y$(否则 $d'_D(x)=0$!)，于是(11)与 $|d'_D(x)|\leqslant1$ 一起推出

$|d'_D(x)|=1, d'_D(x)=(x-y)/|x-y|$. 其余结论是明显的. □

**4.4.10 定理** 设 $x\in\overline{D}\subset\mathrm{R}^n$, 则 $\partial d_D(x)=\mathrm{co}(Z\cup\{0\})$, 其中

$$Z=\left\{\lim_K\frac{z_k}{|z_k|}: z_k\perp_{x_k}D, x_k\to x, z_k\to 0\right\}.$$

**证** 由 4.4.7 与 4.4.9 有 $\partial d_D(x)\subset\mathrm{co}(Z\cup\{0\})$. 显然 $0\in\partial d_D(x)$, 只要证 $Z\subset\partial d_D(x)$. 设 $z=\overline{\lim_K}z_k/|z_k|, z_k\perp x_kD, x_k\to x$, $z_k=y_k-x_k\to 0, d_D(y_k)=|y_k-x_k|$, 则 $|z_k|=d_D(y_k)-d_D(x_k)=\langle u_k, y_k-x_k\rangle, u_k\in\partial d_D(x'_k), x'_k\in[x_k, y_k]$(4.3.5). 因 $|u_k|\leqslant 1$, 故必 $u_k=z_k/|z_k|\to z, z\in\partial d_D(x)$(4.3.2(ii)).

## §5 极大函数

对于由"有限构成过程"作出的函数, 如有限和、有限积及复合函数, 上节给出了适当的"次微分规则". 对于由一"无限构成过程"作出的函数, 如无限和、极限函数、上确界函数等, 亦需类似的规则. 本节对所谓极大函数解决这一问题, 其结果在非光滑最优化理论中有重要应用.

设 $T$ 是一紧度量空间, $\Omega\subset X$ 是一开集, 给定函数族 $\{f_t\in C(\Omega): t\in T\}$, 考虑"极大函数" $f(x)=\max_T f_t(x)$. 有时表 $f_t(x)$ 为 $f(t,x)$ 是方便的. 假定 $f(\cdot, x)$ 为 usc 而 $\mathrm{Lip} f_t\leqslant k<\infty$, 于是 $f:\Omega\to\mathrm{R}$ 有定义且 $\mathrm{Lip} f\leqslant k$. 约定 $T_x=T(x)=\{t: f_t(x)=f(x)\}$;

$$\partial_T f_t(x)=\overline{\mathrm{co}}\{u: u\text{ 是}\{u_n\}\text{ 的弱}^*\text{极限点},\quad u_n\in\partial f_{t_n}(x_n), t_n\to t, x_n\to x\},\tag{1}$$

其中$\overline{\mathrm{co}}$记弱$^*$闭包, 本节中概如此.

**4.5.1 定理**(Clarke, 1973) 任给 $x\in\Omega$, 有

$$\partial f(x)\subset\overline{\mathrm{co}}\bigcup_{t\in T_x}\partial_T f_t(x).\tag{2}$$

若 $\forall\, t\in T_x: f_t$ 在 $x$ 正则, $\partial_T f_t(x)=\partial f_t(x)$, 则 $f$ 在 $x$ 正则且

$$\partial f(x)=\overline{\mathrm{co}}\bigcup_{t\in T_x}\partial f_t(x)$$

$$=\left\{\int_T u_t \mathrm{d}\mu(t);\mu\in P(T_x),u_t\in\partial f_t(x)\right\}. \tag{3}$$

其中 $P(T_x)$ 记 $T$ 上由 $T_x$ 支承的概率 Radon 测度之全体，积分 $\int_T u_t \mathrm{d}\mu(t)$ 在弱$^*$收敛意义下存在.

**证** 首先证(2)式. 令 $Q=\bigcup_{t\in T_x}\partial_T f_t(x)$，只需证 $f^\circ(x,\cdot)\leqslant s_Q$. 取定 $z\in X$. 取 $x_n\to x,\tau_n\downarrow 0$，使 $\alpha_n\triangleq\tau_n^{-1}\Delta f(x_n,\tau_n z)\to f^\circ(x,z)$. 设 $t_n\in T(x_n+\tau_n z)$，则 $\alpha_n\leqslant\tau_n^{-1}\Delta f_{t_n}(x_n,\tau_n z)=u_n(z),u_n\in\partial f_{t_n}(y_n),y_n\in[x_n,x_n+\tau_n z]$(4.3.5). 可设 $t_n\to t\in T,\{u_n\}$ 有弱$^*$极限点 $u$. 因 $y_n\to x$，由(1)有 $u\in\partial_T f_t(x)$；由 $\alpha_n\leqslant u_n(z)$ 推出 $f^\circ(x,z)\leqslant u(z)$. 由

$$\begin{aligned}f_t(x)&\geqslant\overline{\lim_n}f_{t_n}(x)\\&\geqslant\overline{\lim_n}[f(x_n+\tau_n z)-k|x_n+\tau_n z-x|]\\&=f(x)\end{aligned}$$

得 $f_t(x)=f(x)$，即 $t\in T_x$. 于是 $u\in Q,f^\circ(x,z)\leqslant s_Q(z)$.

其次在附加条件下证(3)式. $\forall\, t\in T_x$，有

$$f^\circ_t(x,z)=\lim_{\tau\downarrow 0}\frac{1}{\tau}\Delta f_t(x,\tau z)\leqslant\underline{\lim_{\tau\downarrow 0}}\frac{1}{\tau}\Delta f(x,\tau z).$$

于是

$$\begin{aligned}f^\circ(x,z)\leqslant s_Q(z)&=\sup_{t\in T_x}\max\langle\partial f_t(x),z\rangle\\&=\sup_{t\in T_x}f^\circ_t(x,z)\leqslant\underline{\lim_{\tau\downarrow 0}}\Delta f(x,\tau z),\end{aligned}$$

这推出 $f$ 在 $x$ 正则. 由 Aubin 的一个结果可推知(3)中第二个等号成立(参考[36;p.89]). 在(3)右端任取一元 $u=\int_T u_t\mathrm{d}\mu(t),u_t\in\partial f_t(x),\mu\in P(T_x)$，有

$$\begin{aligned}u(z)&=\int_T u_t(z)\mathrm{d}\mu(t)\leqslant\int_T f^\circ_t(x,z)\mathrm{d}\mu(t)\\&=\int_T\overline{\lim_{\tau\downarrow 0}}\frac{1}{\tau}\Delta f_t(x,\tau z)\mathrm{d}\mu(t)\end{aligned}$$

$$=\overline{\lim_{\tau\downarrow 0}}\frac{1}{\tau}\int_T \Delta f_t(x,\tau z)\mathrm{d}\mu(t)$$

$$\leqslant\overline{\lim_{\tau\downarrow 0}}\frac{1}{\tau}\int_T \Delta f(x,\tau z)\mathrm{d}\mu(t)=f^\circ(x,z)$$

(其中用了控制收敛定理),这得出 $u\in\partial f(x)$. 因此(3)成立. □

在 4.5.1 的应用中,"$\partial_T f_t(x)=\partial f_t(x)$"成立是重要的. 以下推论皆属于这种情况.

**4.5.2 推论** 若 $f(t,x)$对 $t$ 连续,对 $x$ 凸,则 $f(x)$凸,且(3)成立.

**证** 直接看出 $f(x)$凸. 只需说明对 $t\in T_x$ 有 $\partial_T f_t(x)\subset\partial f_t(x)$. 设 $u_n\in\partial f_{t_n}(x_n)$, $t_n\to t$, $x_n\to x$, $u$ 是$\{u_n\}$的弱$^*$极限点. 为记号简便,不妨设 $u_n\xrightarrow{*}u$. 于是

$$u(z)=\overline{\lim_n}u_n(z)\leqslant\overline{\lim_n}\Delta f_{t_n}(x_n,z)$$

$$\leqslant\overline{\lim_n}\left[\Delta f_{t_n}(x,z)+2k|x_n-x|\right]=\Delta f_t(x,z),$$

这表明 $u\in\partial f_t(z)$. □

**4.5.3 推论** 若 $f_t$ 皆严格可微且$(t,x)\mapsto Df_t(x)$连续,则 $f$ 正则,且

$$\partial f(x)=\overline{\mathrm{co}}\{Df_t(x):t\in T_x\}$$

$$=\left\{\int_T Df_t(x)\mathrm{d}\mu(t):\mu\in P(T_x)\right\}(x\in\Omega).\qquad(4)$$

事实上,在所述条件下显然有 $\partial_T f_t(x)=\partial f_t(x)=Df_t(x)$.

**4.5.4 推论** 设 $f_i\in C^{1-0}(\Omega)(1\leqslant i\leqslant n)$, $f=\max_i f_i$,则

$$\partial f(x)\subset\mathrm{co}\bigcup_{i\in I(x)}\partial f_i(x),\ I(x)=\{i:f_i(x)=f(x)\}.\qquad(5)$$

若 $i\in I(x)\Rightarrow f_i$ 在 $x$ 正则,则 $f$ 在 $x$ 正则且(5)为等式.

这是 4.5.1 的直接推论.

下面考虑 4.5.1 的一个变种.

**4.5.5 定理**(Clarke, 1976) 设 $T$ 是任一集, $\Omega\subset\mathrm{R}^n$ 开, $f_t\in C(\Omega)$, $\mathrm{Lip}f_t\leqslant k<\infty$, $f(x)=\sup_{t\in T}f_t(x)$有限, $N\subset\Omega$ 是一

(Lebesgue)零测集，则

$$\partial f(x) \subset \mathrm{co}\{\lim_K f'_{t_k}(x_k): x_k \to x, x_k \overline{\in} N, f_{t_k}(x) \to f(x)\}. \tag{6}$$

**证** 以 $Q$ 记(6)之右端，易见 $Q$ 非空紧凸，故只要证 $f^\circ(x,\cdot) \leqslant s_Q$. 取定 $z \in \mathrm{R}^n \setminus \{0\}$，不妨设 $|z|=1$，令 $\alpha = s_Q(z)$. $\forall\ \varepsilon > 0$，今证 $f^\circ(x,z) \leqslant \alpha + \varepsilon$. 以 $N_t$ 记 $f_t$ 的不可微点集. 取 $\delta \in (0,1)$，使当 $y \in B_{2\delta}(x) \setminus (N \cup N_t)$，$f_t(x) \geqslant f(x) - \delta$ 时 $\langle f'_t(y), z\rangle < \alpha + \varepsilon$. 取定 $t \in T$ 使 $f_t(x) \geqslant f(x) - \delta$. 对几乎所有 $y \in B_\delta(x)$，$[y, y+\delta z] \cap (N \cup N_t)$是 1 维零测集；对这样的 $y$ 及任何 $\lambda \in (0,\delta)$有

$$\Delta f_t(y,\lambda z) = \int_0^\lambda \langle f'_t(y+\tau z), z\rangle \mathrm{d}\tau \leqslant \lambda(\alpha+\varepsilon).$$

由连续性得$\forall\ y \in B_\delta(x)$：$\Delta f_t(y,\lambda z) \leqslant \lambda(\alpha+\varepsilon)$. 若预先取 $t$ 使 $f_t(y+\lambda z) \geqslant f(y+\lambda z) - \lambda^2$，则当 $\lambda$ 与 $|x-y|$ 充分小时可验知 $f_t(x) \geqslant f(x) - \delta$，于是

$$\Delta f(y,\lambda z) \leqslant \Delta f_t(y,\lambda z) + \lambda^2 \leqslant \lambda(\alpha+\varepsilon) + \lambda^2,$$

这推出 $f^\circ(x,z) \leqslant \alpha + \varepsilon$. □

## §6 切　　锥

在非光滑分析与最优化理论中，"切锥"概念有基本的重要性，其意义可与"光滑分析"中的切平面(或切线)概念相对比.

设 $x \in D \subset X$. 约定 $x_k \xrightarrow{D} x \Leftrightarrow x_k \in D$ 且 $x_k \to x$；"当 $k \to \infty$时 $x_k \in D$"意味着对充分大的 $k$ 有 $x_k \in D$. 令

$$H_D(x) = \{z: \forall\ x_k \xrightarrow{D} x, \forall\ t_k \downarrow 0, \forall\ z_k \to z, \text{当 } k\to\infty \text{ 时 } x_k + t_k z_k \in D\}; \tag{1}$$

$$F_D(x) = \{z: \forall\ t_k \downarrow 0, \forall\ z_k \to z, \text{当 } k\to\infty \text{ 时 } x + t_k z_k \in D\}; \tag{2}$$

$$T_D(x) = \{z: \forall\ x_k \xrightarrow{D} x, \forall\ t_k \downarrow 0, \exists\ z_k \to z:$$

$$x_k + t_k z_k \in D\}; \tag{3}$$

$$K_D(x) = \{z: \qquad \exists\, t_k \downarrow 0, \exists\, z_k \to z: x + t_k z_k \in D\}. \tag{4}$$

不难直接验证 $H_D(x)$ 与 $F_D(x)$ 是开锥，$T_D(x)$ 与 $K_D(x)$ 是闭锥. 文献中通常称 $H_D(x)$ 为 $D$ 在 $x$ 的超切锥，称 $T_D(x)$ 为 Clarke 切锥. 关于锥(1)～(4)及其它诸如此类的锥的名称与记号颇为杂乱. 一般地，若以 $A_D(x)$ 表 $D$ 在 $x$ 的某种"切锥"，则 $A_D(\cdot): D \to 2^X$ 是一"锥值"映射，亦记作 $A(D,\cdot)$，通常要求它满足：$A(D,x)=A(D-x,0)$；$A(D\cap B_\delta(0),x)=A(D,x)$ ($\forall\ \delta>0$)；$x\in D^\circ \Rightarrow A(D,x)=X$. 以上性质显然为(1)～(4)所具有. 约定 $A_D^*(x)=(A_D(x))^*$；称 $N_D(x)\triangleq -T_D^*(x)$ 为 $D$ 在 $x$ 的法锥. 若 $T_D(x)=K_D(x)$，则说 $D$ 在 $x$ 正则.

不难验证以下包含关系：

$$H_D(x) \subset F_D(x) \cap T_D(x) \subset F_D(x) \cup T_D(x) \subset K_D(x). \tag{5}$$

(5)中的包含都不能换成等式，以次例说明之.

**4.6.1 例** 设 $f(x)=x\sin x^{-1}$，$f(0)=0$，$D=\mathrm{epi} f\subset \mathbb{R}^2$，则可依(1)～(4)直接看出 $H_D(0)=\varnothing$；$F_D(0)=\{x\in\mathbb{R}^2: |x_1|<x_2\}$；$T_D(0)=\{0\}\times\mathbb{R}_+$；$K_D(0)=\{x\in\mathbb{R}^2: |x_1|\geqslant -x_2\}$.

关于切锥的一些简单性质汇集于下.

**4.6.2 命题** 设 $(x,y)\in D\times E\subset X\times Y$. (i) $F_D(x)\subset \mathrm{cone}(D^\circ - x)$；当 $D$ 凸时等式成立. (ii) 若 $D=\bigcap_1^n D_i$，则 $F(D,x)=\bigcap F(D_i,x)$；$F(D_1,x)\cap K(D_2,x)\subset K(D_1\cap D_2,x)$；$H(D_1,x)\cap T(D_2,x)\subset T(D_1\cap D_2,x)$. (iii) 对 $A=H,F,T,N$ 有 $A_{D\times E}(x,y)=A_D(x)\times A_D(y)$. (iv) 若 $D$ 在 $x$ 或 $E$ 在 $y$ 正则，则 $K_{D\times E}(x,y)=K_D(x)\times K_E(y)$.

**证** 只证(i)，其余是明显的. 易见 $F_D(x)\subset\mathrm{cone}(D^\circ - x)$. 若 $D$ 凸，$z=\lambda(y-x)$，$\lambda>0$，$y\in D^\circ$，则当 $t\downarrow 0$，$w\to z$ 时

$$x + tw = (1-\lambda t)x + \lambda t(y + \lambda^{-1}(w-z)) \in D,$$

可见 $z\in F_D(x)$，这得出 $F_D(x)=\mathrm{cone}(D^\circ - x)$. □

给定 $f:X\to\mathrm{R}$,"水平集"$D=\{x:f(x)\leqslant f(\bar{x})\}$在 $\bar{x}$ 的切锥可能对 $f$ 在 $\bar{x}$ 邻近的局部性质提供重要信息,因此特别值得考察.

**4.6.3 定理** 设 $f$ 在 $\bar{x}$ 邻近为 Lip,$0\bar{\in}\partial f(\bar{x})$,$K_0=\{z:f^\circ(\bar{x},z)<0\}$,则

$$\overline{K_0}=\{z:f^\circ(\bar{x},z)\leqslant 0\};\tag{6}$$

$$K_0\subset H_D(\bar{x})\subset F_D(\bar{x}),\overline{K_0}\subset T_D(\bar{x})\subset K_D(\bar{x});\tag{7}$$

$$\begin{aligned}\mathrm{R}_-\,\partial f(\bar{x})\supset H_D^*(\bar{x})\supset F_D^*(\bar{x}),\mathrm{R}_+\,\partial f(x)\\ \supset N_D(\bar{x})\supset -K_D^*(\bar{x}).\end{aligned}\tag{8}$$

若 $f$ 在 $\bar{x}$ 正则,则(7)(8)为等式,$D$ 在 $\bar{x}$ 正则.

**证** 必定 $K_0\neq\varnothing$($f^\circ(\bar{x},\cdot)\geqslant 0\Rightarrow 0\in\partial f(\bar{x})$!). 取 $z_0\in K_0$. 若 $f^\circ(x,z)\leqslant 0$,则$\forall\ t\in(0,1):z_t\triangleq tz_0+t'z\in K_0,z_t\to z(t\downarrow 0)$,因此 $z\in\overline{K_0}$. 这得出(6). 若 $x_k\xrightarrow{D}\bar{x},t_k\downarrow 0,z_k\to z\in K_0$,则

$$\begin{aligned}f(x_k+t_kz_k)-f(\bar{x})\leqslant&\Delta f(x_k,t_kz)+|f(x_k+t_kz_k)\\&-f(x_k+t_kz)|\\ \leqslant&t_kf^\circ(\bar{x},z)+o(t_k),\end{aligned}$$

这推出 $k\to\infty$ 时 $x_k+t_kz_k\in D$,因此 $z\in H_D(\bar{x})$. 这得出 $K_0\subset H_D(\bar{x})$;结合(5)得出(7).

因 $\partial f(\bar{x})$弱$^*$紧而 $0\bar{\in}\partial f(\bar{x})$,用一标准的 Moore-Smith 极限论证可指明 $\mathrm{R}.\partial f(\bar{x})$弱$^*$闭. 若 $u\in X^*\backslash\mathrm{R}.\partial f(\bar{x})$,则由分离定理有 $z\in X:u(z)<\langle\mathrm{R}.\partial f(\bar{x}),z\rangle$;这推出 $u(z)<0,f^\circ(\bar{x},z)\leqslant 0$,从而 $u\bar{\in}\overline{K_0}^*=K_0^*$. 因此 $K_0^*\subset\mathrm{R}_-\partial f(\bar{x})$;显然 $K_0^*\supset\mathrm{R}_-\partial f(\bar{x})$,故得 $K_0^*=R_-\partial f(\bar{x})$. 这结合(7)得(8).

下面设 $f$ 在 $\bar{x}$ 正则. 若 $z\in K_D(\bar{x})$,则有 $t_k\downarrow 0,z_k\to z;x_k\triangleq\bar{x}+t_kz_k\in D$,即 $f(x_k)\leqslant f(\bar{x})$,于是

$$\begin{aligned}f^\circ(\bar{x},z)=&\lim_k t_k^{-1}\Delta f(\bar{x},t_kz)\\ \leqslant&\lim_k t_k^{-1}[f(\bar{x}+t_kz)-f(x_k)]=0,\end{aligned}$$

故 $z\in\overline{K_0}$. 这得 $F_D(\bar{x})\subset K_D(\bar{x})\subset\overline{K_0}$. $\forall\ z\in F_D(\bar{x})$,若 $z\bar{\in}K_0$,则必

$f^\circ(\bar{x},z)=0$，于是$\exists\ u\in\partial f(\bar{x})$：$u(z)=0$. 因 $u\neq 0$，故$\exists\ z_k\to z$：$u(z_k)>0$，当 $t\downarrow 0,k\to\infty$时 $\bar{x}+tz_k\in D$，于是

$$0<u(z_k)\leqslant f^\circ(\bar{x},z_k)=\lim_{t\downarrow 0}\frac{1}{t}\Delta f(\bar{x},tz_k)\leqslant 0,$$

得出矛盾. 因此 $z\in K_0$，故得 $F_D(\bar{x})\subset K_0$. 由此可见(7)(从而(8))为等式. □

**4.6.4 推论** 设 $f$ 在 $\bar{x}$ 邻近为 $C^1$(或可微凸)函数，$f'(\bar{x})\neq 0$，$K_0=\{z:\langle f'(\bar{x}),z\rangle<0\}$，则$\overline{K_0}=\{z:\langle f'(\bar{x}),z\rangle\leqslant 0\}=T_D(\bar{x})=K_D(\bar{x})$，$K_0=H_D(\bar{x})=F_D(\bar{x})$，$K_D^*(\bar{x})=F_D^*(\bar{x})=\mathrm{R}_-f'(\bar{x})$.

**注** 若仅假定 $f$ 在 $\bar{x}$ 可微且 $f'(\bar{x})\neq 0$，则可直接证明 $F_D(\bar{x})\subset K_0$，$K_D(\bar{x})\subset\overline{K_0}$.

若令 $\varphi(x,r)=f(x)-r$，则 $\mathrm{epi}f=\{(x,r):\varphi(x,r)\leqslant 0\}$，$\varphi^\circ((x,r),(z,s))=f^\circ(x,z)-s$，$\partial\varphi(x,f(x))=\partial f(x)\times\{-1\}$；$f$ 在 $x$ 正则$\Rightarrow\varphi$ 在$(x,f(x))$正则. 于是可用 4.6.3 得出：

**4.6.5 推论** 设 $f$ 在 $x$ 邻近为 Lip，则

$$T_{\mathrm{epi}f}(x,f(x))=\mathrm{epi}f^\circ(x,\cdot);\tag{9}$$

$$N_{\mathrm{epi}f}(x,f(x))=\mathrm{R}_+(\partial f(x)\times\{-1\});\tag{10}$$

$$\partial f(x)=\{u\in X^*:(u,-1)\in N_{\mathrm{epi}f}(x,f(x))\}.\tag{11}$$

若 $f$ 在 $x$ 正则，则 $\mathrm{epi}f$ 在$(x,f(x))$正则.

**证** 由 4.6.3 有 $\mathrm{epi}f^\circ(x,\cdot)\subset T_{\mathrm{epi}f}(x,f(x))$. 若$(z,r)\in T_{\mathrm{epi}f}(x,f(x))$，则$\forall\ x_k\to x,t_k\downarrow 0$，$\exists\ (z_k,r_k)\to(z,r)$：$(x_k,f(x_k))+t_k(z_k,r_k)\in\mathrm{epi}f$，即 $f(x_k+t_kz_k)\leqslant f(x_k)+t_kr_k$，于是

$$\begin{aligned}&\varlimsup_k t_k^{-1}\Delta f(x_k,t_kz)\\&\leqslant\varlimsup_k t_k^{-1}[f(x_k+t_kz)-f(x_k+t_kz_k)+t_kr_k]\leqslant r,\end{aligned}$$

这推出 $f^\circ(x,z)\leqslant r$. 因此(8)式成立，余下的结论是明显的. □

**注** (11)表明 $\partial f(x)$由 $T_{\mathrm{epi}f}(x,f(x))$唯一确定.

**4.6.6 例** 设 $f(x)=|x|\,(x\in X)$，则 $\partial f(0)=\bar{B}_1(0)\subset X^*$ (4.2.2)，$f^\circ(0,z)=\max\langle\partial f(0),z\rangle=|z|=f(z)$，于是由(9)(10)

有

$$T_{\mathrm{epi}f}(0,0)=\mathrm{epi}f^{\circ}(0,\cdot)=\mathrm{epi}f,$$

$$N_{\mathrm{epi}f}(0,0)=\{(u,-r)\in X^{*}\times\mathrm{R}:|u|\leqslant r\}.$$

再回到一般的点集 $D\subset X$. 下面指明 $T_D(x)$ 与 $\partial d_D(x)$ 有密切联系.

**4.6.7 命题** 设 $x\in D\subset X$, 则

$$T_D(x)=\{z:d^{\circ}_D(x,z)=0\}\subset\{z:d_D(x+tz)=0(t)(t\downarrow 0)\};\tag{12}$$

$$N_D(x)\supset\mathrm{R}_+\,\partial d_D(x)\supset\partial\delta_D(x);\tag{13}$$

$$\partial d_D(x)\subset N_D(x)\cap\overline{B}_1(0).\tag{14}$$

若 $D$ 凸, 则(12)~(14)中的包含是等式.

**证** 首先设 $z\in T_D(x)$. $\forall\, t_k\downarrow 0, x_k\to x$, 取 $y_k\in D$, 使 $|x_k-y_k|=d_D(x_k)+0(t_k)$, 则 $y_k\xrightarrow{D}x$, 于是 $\exists\, z_k\to z: y_k+t_kz_k\in D$,

$$\overline{\lim_k}\,t_k^{-1}\Delta d_D(x_k,t_kz)\leqslant\overline{\lim_k}\,t_k^{-1}[|x_k+t_kz-y_k-t_kz_k|-d_D(x_k)]\leqslant 0,$$

这推出 $d^{\circ}_D(x,z)=0$, 更有 $d_D(x+tz)=0(t)\,(t\downarrow 0)$. 反之, 若 $d^{\circ}_D(x,z)=0$, 则 $\forall\, x_k\xrightarrow{D}x, t_k\downarrow 0$, 有 $d_D(x_k+t_kz)=0(t_k)$. 取 $y_k\in D$, 使 $|x_k+t_kz-y_k|=0(t_k)$, 则 $z_k\triangleq t_k^{-1}(y_k-x_k)\to z$, $x_k+t_kz_k=y_k\in D$, 可见 $z\in T_D(x)$. 于是(12)得证.

由(12)推出 $\langle\partial d_D(x),T_D(x)\rangle\leqslant 0$, 因此 $\mathrm{R}_+\partial d_D(x)\subset N_D(x)$. 若 $u\in\partial\delta_D(x)=(x-D)^{*}$, 则 $x$ 是 $-u|D$ 的极小点, 于是 $0\in -u+|u|\partial d_D(x)$(4.3.7), 从而 $u\in\mathrm{R}_+\partial d_D(x)$. 这证得(13). 由(13)与 $\mathrm{Lip}d_D\leqslant 1$ 得出(14).

下面设 $D$ 凸, 从而 $d_D$ 凸. 因 $d_D$ 在 $x$ 正则, 故(12)为等式. 任给 $u\in N_D(x)$, $y\in D$. 若 $x_k\xrightarrow{D}x, t_k\downarrow 0$, 则 $y-x_k\to y-x$; 由 $D$ 凸有 $x_k+t_k(y-x_k)\in D$, 于是 $y-x\in T_D(x)$. 因此 $\langle u,y-x\rangle\leqslant 0$. 这证得 $\langle N_D(x),D-x\rangle\leqslant 0$, 即 $N_D(\bar{x})\subset(x-D)^{*}=\partial\delta_D(x)$, 故(13)

是等式. 若 $u\in N_D(x)\cap\overline{B}_1(0)$，则 $\langle u,D\rangle\leqslant u(x)$，于是

$$u(z)\leqslant\inf_{y\in D}[u(y)+u(z-y)]\leqslant u(x)+d_D(z)(\forall z\in X),$$

这得出 $u\in\partial d_D(x)$. 故(14)是等式. □

最后指出，在 $H_D(x)$ 与 $T_D(x)$ 之间，存在如下述定理表达的准确关系.

**4.6.8 定理**(Rockafellar, 1980) 若 $H_D(x)\neq\varnothing$，则 $H_D(x)=\operatorname{int}T_D(x)$；当 $D\subset\mathbf{R}^n$ 闭时不必假定 $H_D(x)\neq\varnothing$.

**证** 由(5)，只需证 $\operatorname{int}T_D(x)\subset H_D(x)$. $\forall y\in\operatorname{int}T_D(x)$，取 $z\in H_D(x)$（$H_D(x)\neq\varnothing$ 用于此！），当 $t\downarrow 0$ 时有

$$y=(y-tz)+tz\in T_D(x)+H_D(x).$$

可见 $\operatorname{int}T_D(x)\subset T_D(x)+H_D(x)$，于是只需证 $T_D(x)+H_D(x)\subset H_D(x)$. 设 $y\in T_D(x)$，$z\in H_D(x)$. $\forall x_k\xrightarrow{D}x$，$t_k\downarrow 0$，$w_k\to y+z$，由 $y\in T_D(x)$ 有 $y_k\to y$：$x_k+t_ky_k\in D$；由 $x_k+t_ky_k\xrightarrow{D}x$，$w_k-y_k\to z$ 与 $z\in H_D(x)$ 得出，当 $k\to\infty$ 时 $(x_k+t_ky_k)+t_k(w_k-y_k)=x_k+t_kw_k\in D$，可见 $y+z\in H_D(x)$.

若 $D\subset\mathbf{R}^n$ 闭，$z\in\operatorname{int}T_D(x)$，则 $\langle N_D(x)\backslash\{0\},z\rangle<0$(2.2.2)，这结合(13)得 $\max\langle S^{n-1}\cap\partial d_D(x),z\rangle<0$；于是当 $y$ 邻近 $x$ 且 $d'_D(y)$ 存在且非零时 $\langle d'_D(y),z\rangle<0$(4.4.7, 4.4.9). 基于此，用一个类似于 4.4.7 之证的论证得出：对几乎所有邻近 $x$ 的 $y\in D$，当 $t\downarrow 0$，$w\to z$ 时有

$$d_D(y+tw)=\int_0^t\langle d'_D(y+sw),w\rangle\mathrm{d}s\leqslant 0,$$

即 $y+tw\in D$. 由 $D$ 闭，这得出 $z\in H_D(x)$. □

## §7 非 Lipschitz 函数的次微分

将非光滑分析限制于 Lip 函数显然不能满足实际需要，即使 $\delta_D(D\subset X)$ 这样简单的函数在 $\partial D$ 邻近也不是局部 Lip. 有许多人

致力于建立“非 Lipschitz 广义微分学”，主要由于 Rockafellar，Hiriart-Urruty 与 Aubin 等人在 1980 年前后的工作，这一方向取得了一定进展. 关键性的步骤是 Rockafellar 给出了非 Lip 函数广义方向导数的适当定义. 由于条件十分宽泛，要达到多少类似于 Clarke 次微分理论的结果，不可避免要用到某些复杂的构造.

以下设 $f:X\to\overline{\mathbf{R}}, x\in D_f$. 采用 Rockafellar 的记号：

$$(y,\lambda)\downarrow_f x\Leftrightarrow(y,\lambda)\to(x,f(x)),f(y)\leqslant\lambda. \tag{1}$$

**4.7.1 定义** 任给 $z\in X$，称

$$f^\circ(x,z)\triangleq\sup_{\epsilon>0}\ \overline{\lim_{(y,\lambda)\downarrow_f x,t\downarrow 0}}\ \inf_{|w-z|<\epsilon}\frac{1}{t}[f(y+tw)-\lambda] \tag{2}$$

为 $f$ 在 $x$ 沿 $z$ 的广义方向导数；称

$$\partial f(x)\triangleq\{u\in X^*;u\leqslant f^\circ(x,\cdot)\} \tag{3}$$

为 $f$ 在 $x$ 的次微分. 若 $\mathrm{epi}f$ 在 $(x,f(x))$ 正则，则说 $f$ 在 $x$ 是正则的.

与 4.3.1 比较，$\partial f(x)$ 的定义式并无改变，但 $f^\circ(x,z)$ 的定义显著地复杂了，需要某些说明.

1° $f^\circ(x,\cdot):X\to\mathbf{R}\cup\{\pm\infty\}$ 显然是正齐次的；较精细的分析可指明它亦次可加，从而是凸的. 显然 $\partial f(x)\neq\varnothing\Rightarrow f^\circ(x,\cdot)>-\infty$；$\partial f(x)$ 弱*闭凸（但未必弱*紧）.

2°若 $f$ 在 $x$ 为 lsc，则(2)可略作简化：

$$f^\circ(x,z)=\sup_{\epsilon>0}\ \overline{\lim_{\substack{(y,f(y))\to(x,f(x))\\ t\downarrow 0}}}\ \inf_{|w-z|<\epsilon}\frac{1}{t}\Delta f(y,tw). \tag{4}$$

3°设 $f$ 在 $x$ 邻近为 Lip，则(4)与 §3(1)一致，从而(3)所定义的 $\partial f(x)$ 就是 Clarke 次微分. 若 $f^\circ(x,\cdot)=D_+f(x,\cdot)$，则由 4.6.5 知 $\mathrm{epi}f$ 在 $(x,f(x))$ 正则. 反之，若 $\mathrm{epi}f$ 在 $(x,f(x))$ 正则，则必 $f^\circ(x,\cdot)=D_+f(x,\cdot)$. 事实上，若 $\tau_k\downarrow 0, s_k\triangleq\tau_k^{-1}\Delta f(x,\tau_k z)\to s$，则由 $(x,f(x))+\tau_k(z,s_k)\in\mathrm{epi}f$ 推出 $(z,s)\in K_{\mathrm{epi}f}(x,f(x))=T_{\mathrm{epi}f}(x,f(x))$；于是 $\forall\ x_k\to x,t_k\downarrow 0,\exists\ (z_k,r_k)\to(z,s)$，使得 $(x_k,f(x_k))+t_k(z_k,r_k)\in\mathrm{epi}f$，即 $f(x_k+t_kz_k)\leqslant f(x_k)+t_kr_k$，因此

$$\overline{\lim_k} t_k^{-1}\Delta f(x_k, t_k z) \leqslant \overline{\lim_k} t_k^{-1}[\,|f(x_k + t_k z) - f(x_k + t_k z_k)| + t_k r_k] \leqslant s.$$

这推出 $f^\circ(x,z)=s=D_+f(x,z)$. 因此，$f$ 依 4.7.1 在 $x$ 正则，恰如它依 4.3.1 在 $x$ 正则.

**4.7.2 例** 设 $x\in D\subset X$，则对指示函数 $\delta_D$ 有

$$\delta^\circ_D(x,z) = \sup_{\varepsilon>0} \overline{\lim_{\substack{y\in D, y\to x\\ \lambda\downarrow 0, t\downarrow 0}}} \inf_{|w-z|<\varepsilon} \frac{1}{t}[\delta_D(y+tw) - \lambda]. \tag{5}$$

这与 §6(3)对照看出，当 $z\in T_D(x)$ 时 $\delta^\circ_D(x,z)=0$，否则 $\delta^\circ_D(x,z)=\infty$，即 $\delta^\circ_D(x,\cdot)$ 是 $T_D(x)$ 的指示函数. 由此得出 $\partial\delta_D(x)=(0-T_D(x))^*=N_D(x)$. 当 $D$ 凸时这与 4.6.7 的结论一致(注意 4.6.7 中的 $\partial\delta_D(x)$ 依 4.2.1!). 因 $\mathrm{epi}\delta_D=D\times\mathrm{R}_+$，故 $\delta_D$ 在 $x$ 正则 $\Leftrightarrow D\times\mathrm{R}_+$ 在 $(x,0)$ 正则 $\Leftrightarrow T_D(x)\times\mathrm{R}_+=K_D(x)\times\mathrm{R}_+$(用 4.6.2)$\Leftrightarrow$ $D$ 在 $x$ 正则.

**4.7.3 命题** 若 $x\in D_f$，则以下结论成立：

$$T_{\mathrm{epi}f}(x,f(x)) = \mathrm{epi}f^\circ(x,\cdot); \tag{6}$$

$$\partial f(x) = \{u\in X^*: (u,-1)\in N_{\mathrm{epi}f}(x,f(x))\}; \tag{7}$$

$$f^\circ(x,z) = \sup\langle\partial f(x), z\rangle\ (\partial f(x)\neq\varnothing). \tag{8}$$

**证** 若 $(z,r)\in T_{\mathrm{epi}f}(x,f(x))$，则 $\forall\ (x_k,\lambda_k)\downarrow_f x, t_k\downarrow 0, \varepsilon>0$，$\exists\ (z_k,r_k)\to(z,r): (x_k,\lambda_k)+t_k(z_k,r_k)\in\mathrm{epi}f$，即 $f(x_k+t_kz_k)\leqslant\lambda_k+t_kr_k$，于是

$$\overline{\lim_k} \inf_{|w-z|<\varepsilon} \frac{1}{t_k}[f(x_k+t_kw)-\lambda_k]$$

$$\leqslant \overline{\lim_k} \frac{1}{t_k}[f(x_k+t_kz_k)-\lambda_k] \leqslant r,$$

这推出 $f^\circ(x,z)\leqslant r$. 反之，若 $(z,r)\overline{\in} T_{\mathrm{epi}f}(x,f(x))$，则 $\exists\ (x_k,\lambda_k)\downarrow_f x, t_k\downarrow 0, \varepsilon>0, \forall\ (w,s)\in B_\varepsilon(z)\times[r-\varepsilon,r+\varepsilon]: (x_k,\lambda_k)+t_k(w,s)\overline{\in}\mathrm{epi}f$，即 $f(x_k+t_kw)>\lambda_k+t_ks$，于是

$$\overline{\lim_k} \inf_{|w-z|<\varepsilon} \frac{1}{t_k}[f(x_k+t_kw)-\lambda_k] \geqslant r+\varepsilon,$$

这推出 $f^\circ(x,z)\geqslant r+\varepsilon>r$. 这就证得(6)式.

$\forall u\in X^*$,利用已证之(6),有 $u\in\partial f(x)\Leftrightarrow u\leqslant f^\circ(x,\cdot)\Leftrightarrow$ $\forall (z,r)\in T_{\mathrm{epi}f}(x,f(x)):u(z)\leqslant r\Leftrightarrow(u,-1)\in -T^*_{\mathrm{epi}f}(x,f(x))=$ $N_{\mathrm{epi}f}(x,f(x))$,这得出(7).

由(3)直接得 $f^\circ(x,z)\geqslant\sup\langle\partial f(x),z\rangle$. 相反不等式之证明类似于 4.3.2(iii). □

**注** (6)(7)(8)可分别与§6中(9)(11)及4.3.2所蕴涵的等式 $f^\circ(x,z)=\max\langle\partial f(x),z\rangle$ 对照. (6)(7)给出了 $f^\circ(x,\cdot)$ 与 $\partial f(x)$ 的几何刻划. 若 $f\in C^1(\mathrm{R})$,则 $(f'(x),-1)$ 是曲线 $y=f(x)$ 的法向量. (7)可看作是上述事实的推广.

为得出较深入的结论,一般广义实泛函显然过宽,而 Lip 函数又嫌过窄. Rockafellar 引入了一个中间类型,它能保持 Lip 函数的某些特点.

**4.7.4 定义** 任给 $z\in X$,令

$$f^+(x,z)=\overline{\lim_{(y,\lambda)\downarrow_f x,t\downarrow 0,w\to z}}\frac{1}{t}[f(y+tw)-\lambda];\tag{9}$$

$$D_f(x)=\{z:f^+(x,z)<\infty\}.\tag{10}$$

若 $z\in D_f(x)$,则说 $f$ 在 $x$ 沿 $z$ 为方向 Lip;若 $D_f(x)\neq\varnothing$,则说 $f$ 在 $x$ 为方向 Lip.

直接看出 $f^\circ(x,z)\leqslant f^+(x,z)$;若 $f$ 在 $x$ 邻近为 Lip,则 $f^\circ(x,z)=f^+(x,z)$. 鉴于(9)比(2)简单,确定是否 $f^\circ(x,z)=f^+(x,z)$ 有其重要性.

**4.7.5 定理** 设 $\pi:X\times\mathrm{R}\to X$ 为投影,则

$$f^+(x,z)=\inf\{r:(z,r)\in H_{\mathrm{epi}f}(x,f(x))\};\tag{11}$$

$$D_f(x)=\pi H_{\mathrm{epi}f}(x,f(x)).\tag{12}$$

若 $D_f(x)\neq\varnothing$,则

$$D_f(x)=\mathrm{int}\{z:f^\circ(x,z)<\infty\};\tag{13}$$

在 $D_f(x)$ 内,$f^\circ(x,z)=f^+(x,z)$ 且 $f^\circ(x,\cdot)$ 连续.

**证** 若 $f^+(x,z)<r$,则 $\forall (x_k,\lambda_k)\downarrow_f x,t_k\downarrow 0,(z_k,r_k)\to(z,$

$r$)，当 $k\to\infty$ 时 $t_k^{-1}[f(x_k+t_kz_k)-\lambda_k]<r_k$，即 $(x_k,\lambda_k)+t_k(z_k,r_k)\in$ epi$f$，可见 $(z,r)\in H_{\text{epi}f}(x,f(x))$. 由此推出(11)之右端 $\leqslant f^+(x,z)$. 类似地可证相反的不等式，因此(11)成立. 由(11)直接得出(12).

下面设 $D_f(x)\neq\varnothing$，于是 $H_{\text{epi}f}(x,f(x))\neq\varnothing$，从而

$$H_{\text{epi}f}(x,f(x)) = \text{int epi}f^\circ(x,\cdot) \tag{14}$$

(4.6.8，4.7.3). 若 $z\in D_f(x)$，则 $\exists\ r\in\mathbf{R}:(z,r)\in\text{int epi}f^\circ(x,\cdot)$，于是当 $w\to z$ 时 $f^\circ(x,w)\leqslant r$，可见 $z$ 属于(13)之右端. 反之，若 $\exists\ \varepsilon>0:f^\circ(x,B_\varepsilon(z))<\infty$，则可设 $\exists\ r\in\mathbf{R}:f^\circ(x,B_\varepsilon(z))<r-1$ (参考 4.1.1，注意(6)推出 $f^\circ(x,\cdot)$ 凸且 lsc)，于是 $(z,r)\in\text{int epi}f^\circ(x,\cdot)$，从而 $z\in D_f(x)$. 故(13)成立. 其次，$\forall\ z\in D_f(x)$，有

$$\begin{aligned} f^\circ(x,z) &=\inf\{r:f^\circ(x,z)\leqslant r\} = \inf\{r:(z,r)\in\text{epi}f^\circ(x,\cdot)\} \\ &=\inf\{r:(z,r)\in H_{\text{epi}f}(x,f(x))\} = f^+(x,z). \end{aligned}$$

$f^\circ(x,\cdot)$ 在 $D_f(x)$ 内的连续性由 4.1.1 推出. □

下面的命题提供了"方向 Lip 函数"的若干例子.

**4.7.6 命题** 以下每个条件蕴涵 $D_f(x)\neq\varnothing$：(i) $f$ 在 $x$ 邻近为 Lip；(ii) $f$ 凸且至少在一点 $\hat{x}\in D_f$ 连续；(iii) $X$ 由闭凸锥 $X_+$ 导入序 $\leqslant$，$X^\circ{}_+\neq\varnothing$，$f$ 关于 $\leqslant$ 单调增；(iv) $f=\delta_D$，$x\in D\subset X$，$H_D(x)\neq\varnothing$；(v) $X=\mathbf{R}^n$，$f$ 在 $x$ 邻近 lsc，$\partial f(x)$ 非空且不含整条直线.

**证** (i)不必考虑. (ii) $\Rightarrow\exists\ y\in X,\delta>0,r\in\mathbf{R}:f(B_\delta(y))\leqslant r-1\Rightarrow(y,r)\in\text{int epi}f$，这结合 epi$f$ 凸得出 $(y-x,r-f(x))\in H_{\text{epi}f}(x,f(x))$. 若(iii)满足，取 $z\in X^\circ{}_+$，则当 $(y,\lambda)\downarrow_f x,t\downarrow 0,w\to -z$ 时 $f(y+tw)-\lambda\leqslant f(y+tw)-f(y)\leqslant 0$，这推出 $f^+(x,-z)\leqslant 0$，$-z\in D_f(x)$. 对于(iv)有 $H_{D\times\mathbf{R}_+}(x,0)=H_D(x)\times(0,\infty)\neq\varnothing$，因而 $D_f(x)\neq\varnothing$(参考 4.6.2，4.7.2). 对于(v)，$\partial f(x)$ 不含整条直接推出 $D\triangleq\{z:f^\circ(x,z)<\infty\}$ 不含于任何真子空间，因此 $D^\circ\neq\varnothing$. 不妨设 epi$f$ 闭，于是存在 $(z,r)\in\text{int epi}f^\circ(x,\cdot)=H_{\text{epi}f}(x,f(x))$ (4.6.8). □

**4.7.7 定义** 任给 $x\in D_f$，称

$$\partial^{\infty} f(x) \triangleq \{u \in X^* : (u,0) \in N_{\mathrm{epi} f}(x, f(x))\} \tag{15}$$

为 $f$ 在 $x$ 的渐近次微分.

直接看出 $\partial^{\infty} f(x)$是弱$^*$闭凸锥,$0 \in \partial^{\infty} f(x)$,且

$$\partial^{\infty} f(x) = \{u \in X^* : f^{\circ}(x,z) < \infty \Rightarrow u(z) \leqslant 0\}. \tag{16}$$

若 $f$ 在 $x$ 邻近为 Lip,则用(16)得出 $\partial^{\infty} f(x) = \{0\}$. 当 $X = \mathrm{R}^n$ 时其逆亦真,如以下结果所指明:

**4.7.8 定理** 设 $X = \mathrm{R}^n$,$f$ 在 $x$ 邻近为 lsc,则以下条件等价:(i)$\partial f(x)$非空有界;(ii)$f$ 在 $x$ 邻近为 Lip;(iii)$\partial^{\infty} f(x) = \{0\}$.

关于 Clarke 次微分计算的某些结果可推广至非 Lip 的情况. 下面以"和公式"为例进行说明.

**4.7.9 定理**(Rockafellar) 设 $f, g: X \to \overline{\mathrm{R}}$, $x \in D_g$, $D_g(x) \cap \{z : f^{\circ}(x,z) < \infty\} \neq \varnothing$,则

$$\partial(f+g)(x) \subset \partial f(x) + \partial g(x), \tag{17}$$

且(17)之右端弱$^*$闭. 若 $f, g$ 在 $x$ 正则,则(17)是等式且当 $\partial f(x)$ 与 $\partial g(x)$非空时 $f+g$ 在 $x$ 正则.

**证** 令 $h = f + g$. 首先证

$$h^{\circ}(x,z) \leqslant f^{\circ}(x,z) + g^{\circ}(x,z) \tag{18}$$

(约定$\infty - \infty = \infty$). 先设 $z \in D_g(x)$, $f^{\circ}(x,z) < \infty$, $g^{\circ}(x,z) < r$. 设 $\lambda = \lambda_1 + \lambda_2$,当$(y,\lambda_2) \downarrow_g x$, $t \downarrow 0$, $w \to z$ 时 $t^{-1}[g(y + \lambda w) - \lambda_2] < r$,于是

$$\begin{aligned}
&\overline{\lim_{(y,\lambda) \downarrow_h x, t \downarrow 0}} \inf_{|w-z|<\varepsilon} \frac{1}{t}[h(y+tw) - \lambda] \\
&\leqslant \overline{\lim_{(y,\lambda_1) \downarrow_f x, t \downarrow 0}} \inf_{|w-z|<\varepsilon} \left\{\frac{1}{t}[f(y+tw) - \lambda_1] + r\right\} \\
&\leqslant f^{\circ}(x,z) + r.
\end{aligned}$$

令 $\varepsilon \downarrow 0$, $r \downarrow g^{\circ}(x,z)$得出(18). 其次设 $f^{\circ}(x,z) < \infty$, $g^{\circ}(x,z) < \infty$. 取 $\hat{z} \in D_g(x)$, $f^{\circ}(x,\hat{z}) < \infty$,令 $z_t = t\hat{z} + t'z\ (0 < t < 1)$,则由 $g^{\circ}(x,\cdot)$凸与(13)有 $z_t \in D_g(x)$, $f^{\circ}(x,z_t) < \infty$,于是 $h^{\circ}(x,z_t) \leqslant f^{\circ}(x,z_t) + g^{\circ}(x,z_t)$,令 $t \downarrow 0$ 得出(18).

若 $f^\circ(x,0)$ 或 $g^\circ(x,0)$ 为 $-\infty$，则 $h^\circ(x,0)=-\infty$，于是(17)两端皆为空集. 若 $f^\circ(x,0)=g^\circ(x,0)=0$，则

$$\begin{aligned}\partial h(x) &= \{u\in X^*; u\leqslant h^\circ(x,\cdot)\}\\ &\subset\{u\in X^*; u\leqslant f^\circ(x,\cdot)+g^\circ(x,\cdot)\}\\ &=\partial(f^\circ(x,\cdot)+g^\circ(x,\cdot))(0)=\partial f(x)+\partial g(x),\end{aligned}$$

最后一步用到 4.2.5. 以上证明也表明 $\partial f(x)+\partial g(x)$ 为弱*闭.

现设 $f,g$ 在 $x$ 正则，不妨设 $f^\circ(x,0)=g^\circ(x,0)=0$. 利用 $K_{\mathrm{epi}f}(x,f(x))=\mathrm{epi}f^\circ(x,\cdot)$ 不难得出

$$f^\circ(x,z)=\lim_{w\to z,t\downarrow 0}\frac{1}{t}\Delta f(x,tw);$$

$g^\circ(x,z)$ 仿此. 于是

$$h^\circ(x,z)\geqslant f^\circ(x,z)+g^\circ(x,z),$$

从而(18)是等式. 考察上段证明得出(17)亦为等式. 利用(18)易验知 $K_{\mathrm{epi}h}(x,h(x))\subset\mathrm{epi}h^\circ(x,\cdot)$，这推出 $h$ 在 $x$ 正则. □

## §8 广义 Jacobian 与隐函数定理

近年来有一些研究致力于建立"向量值非光滑分析"[120, 125, 131]. 在有限维空间中，由于可利用某种类似于 §4(8) 的构造，引进"向量次微分"是一件比较容易的事情.

给定开集 $\Omega\subset\mathrm{R}^n$ 与 $f=(f_i):\Omega\to\mathrm{R}^m$，设 $K=(\sum|\mathrm{Lip}f_i|^2)^{1/2}<\infty$. 以 $N_f$ 记 $f$ 的不可微点之集，则 $\mathrm{mes}N_f=0$（参看 4.4.7），于是可定义

$$\partial f(x)=\mathrm{co}\{\lim_k f'(x_k); x_k\to x, x_k\overline{\in}N_f\}. \tag{1}$$

如同 §4(8)，"$f'(x_k)$ 收敛"已隐含于表达式中. 称(1)定义的 $\partial f(x)$ 为 $f$ 在 $x(\in\Omega)$ 的广义 Jacobian. 等同 $L(\mathrm{R}^n,\mathrm{R}^m)$ 与 $\mathrm{R}^{m\times n}$（$m\times n$ 矩阵之全体），并在 $\mathrm{R}^{m\times n}$ 中使用 Euclid 范数 $|\cdot|$. 由(1)直接看出，$\partial f(x)\subset\mathrm{R}^{m\times n}$ 是非空凸紧集；比较(1)与 §4(8)得出

$$\partial f(x)\subset\partial f_1(x)\times\cdots\times\partial f_m(x), x\in\Omega. \tag{2}$$

以 $B_{m\times n}$ 记 $\mathrm{R}^{m\times n}$ 中的闭单位球. 类似于 4.3.2 有

**4.8.1 命题** $\partial f:\Omega\to 2^{\mathrm{R}^{m\times n}}$ 有性质:(i)$R(\partial f)\subset KB_{m\times n}$;(ii)$\partial f$ 是闭算子(依 3.1.1);(iii)$\partial f$ 为 usc.

**证** 任给 $x\in\Omega\backslash N_f$,有

$$|f'(x)|^2=\sum|f'_i(x)|^2\leqslant\sum|\mathrm{Lip}f_i|^2=K^2,$$

由此得出(i). 直接从定义式(1)看出(ii). (iii)的证明类似于 4.3.2 (iv)之证. □

**注** 任给零测集 $S\subset\mathrm{R}^n$,令

$$\partial_S f(x)=\mathrm{co}\{\lim_k f'(x_k):x_k\to x,x_k\overline{\in}S\cup N_f\}.\tag{3}$$

当 $m>1$ 时是否 $\partial f(x)=\partial_S f(x)$ 颇成问题. 可以确证的是,$\forall z\in\mathrm{R}^n$,$\partial_S f(x)z$ 与 $S$ 无关(参考[36;p.71]). 下面将(有点隐蔽地)利用这一点.

**4.8.2 中值定理** 设 $[x,y]\subset\Omega$,则

$$f(x)-f(y)\in\mathrm{co}\partial f([x,y])(x-y).\tag{4}$$

**证** 若 $[x,y]\cap N_f$ 是 1 维零测集,则

$$f(x)-f(y)=\int_0^1 f'(y+t(x-y))(x-y)\mathrm{d}t$$

$$\in\mathrm{co}\partial f([x,y])(x-y).$$

固定 $x\in\Omega$,几乎所有 $y\in\Omega$(可设 $\Omega$ 凸)有上述性质,因此由 $f$ 连续与 $\partial f$usc 得出所要证. □

**4.8.3 定理(链规则)** 设 $\varphi=g\circ f$,$g:\mathrm{R}^m\to\mathrm{R}$,$x\in\Omega$,$g$ 在 $u=f(x)$ 邻近为 Lip,则

$$\partial\varphi(x)\subset\mathrm{co}(\partial g(u)\circ\partial f(x)).\tag{5}$$

若 $g$ 在 $u$ 严格可微,则

$$\partial\varphi(x)=Dg(u)\circ\partial f(x).\tag{6}$$

**证** $Q\triangleq\partial g(u)\circ\partial f(x)$ 是紧集,今指明 $\varphi^\circ(x,\cdot)\leqslant S_Q$(参照 4.4.1 之证). 取定 $z\in\mathrm{R}^n$. 取 $x_k\to x$,$t_k\downarrow 0$,使 $t_k^{-1}\Delta\varphi(x_k,t_kz)\to\varphi^\circ(x,z)$. 由 4.3.5 与 4.8.2 有 $u_k\in[f(x_k),f(x_k+t_kz)]$,$\xi_k\in\partial g(u_k)$,使得

$$t_k^{-1}\Delta\varphi(x_k,t_kz) = t_k^{-1}\langle\xi_k,\Delta f(x_k,t_kz)\rangle$$
$$\in \mathrm{co}\langle\xi_k,\partial f([x_k,x_k+t_kz])z\rangle. \tag{7}$$
由(7)推出有 $y_k\in[x_k,x_k+t_kz]$, $A_k\in\partial f(y_k)$，使得
$$t_k^{-1}\Delta\varphi(x_k,t_kz)\leqslant\langle\xi_k,A_kz\rangle. \tag{8}$$
显然 $y_k\to x$, $u_k\to u$. 不妨设 $\xi_k\to\xi$, $A_k\to A$，于是 $\xi\in\partial g(u)$, $A\in\partial f(x)$，从而由(8)得 $\varphi^\circ(x,z)\leqslant\langle\xi,Az\rangle\leqslant s_Q(z)$. 这就证得(5).

下面设 $\xi=D_sg(u)$ 存在，令 $Q=\xi\circ\partial f(x)$，今证 $\varphi^\circ(x,\cdot)=S_Q$（这推出(6)）. 任给 $z\in\mathrm{R}^n$. 容易看出
$$\varphi^\circ(x,z)=\overline{\lim}\{\varphi'(y)z:y\to x,y\overline{\in}N_f\cup N_\varphi\};$$
$$s_Q(z)=\overline{\lim}\{\langle\xi,f'(y)z\rangle:y\to x,y\overline{\in}N_f\cup N_\varphi\},$$
故只需证 $\varphi'(y)-\xi\circ f'(y)\to 0(y\to x,y\overline{\in}N_f\cup N_\varphi)$. $\forall\ \varepsilon>0$，取 $\delta>0$，使 $\partial g(fB_\delta(x))\subset B_\varepsilon(\xi)$. 取定 $y\in B_\delta(x)\backslash(N_f\cup N_\varphi)$. $\forall\ z\in\mathrm{R}^n$，当 $t>0$ 充分小时有
$$t^{-1}\Delta\varphi(y,tz)=t^{-1}[g(f(y+tz)-g(f(y))]$$
$$=\langle\xi_t,t^{-1}\Delta f(y,tz)\rangle$$
$$\in\mathrm{co}\langle\xi_t,\partial f([y,y+tz])z\rangle, \tag{9}$$
其中 $\xi_t\in\partial g(u_t)$, $u_t\in[f(y),f(y+tz)]$. (9)推出 $\varphi'(y)z\in Q_yz$, $Q_y=\partial g(f(y))\circ f'(y)$. 因 $Q_y$ 凸而 $z$ 是任意的，故 $\varphi'(y)\in Q_y$，于是 $\varphi'(y)\in B_\varepsilon(\xi)\circ f'(y)\in\overline{B}_{K\varepsilon}(\xi\circ f'(y))$，恰如所要证. □

**4.8.4 推论** 若 $g:\mathrm{R}^m\to\mathrm{R}^p$ 在 $u=f(x)$ 邻近为 Lip，则
$$\partial(g\circ f)(x)\subset\mathrm{co}(\partial g(u)\circ\partial f(x)). \tag{10}$$
若进而假定 $g\in C^1$，则
$$\partial(g\circ f)(x)=g'(u)\circ\partial f(x). \tag{11}$$
**证** $\forall\ \xi\in\mathrm{R}^p=(\mathrm{R}^p)^*$，反复用 4.8.3 得
$$\xi\circ\partial(g\circ f)(x)=\partial(\xi\circ g\circ f)(x)$$
$$\subset\mathrm{co}(\partial(\xi\circ g)(u)\circ\partial f(x))=\xi\circ\mathrm{co}(\partial g(u)\circ\partial f(x)),$$
由此得出(10). 类似地可证(11). □

**4.8.5 反函数定理** 设 $f:\Omega\subset\mathrm{R}^n\to\mathrm{R}^n$ 为 Lip，$x_0\in\Omega$, $\partial f(x_0)$

$\subset GL(\mathrm{R}^n)$. 则 $f$ 在 $x_0$ 处为局部 Lip 同胚，即存在 $x_0$ 的邻域 $U$ 与 $f(x_0)$ 的邻域 $V$，使 $f:U\to V$ 为双射且 $f^{-1}:V\to U$ 为 Lip.

**证** 首先证 $\exists\ \varepsilon,\delta>0$，使得 $\forall\ z\in S\triangleq S^{n-1}$，

$$\exists\ w\in S:\langle w,\partial f(B_\delta(x_0))z\rangle\geqslant\varepsilon. \tag{12}$$

因 $\partial f(x_0)S\subset\mathrm{R}^n\setminus\{0\}$ 紧，故 $2\varepsilon\triangleq d(0,\partial f(x_0)S)>0$. 取 $\eta>0$ 充分小，使 $Q\triangleq\partial f(x_0)+\eta B_{n\times n}$ 满足 $d(0,QS)\geqslant\varepsilon$. 由 $\partial f$usc，$\exists\ \delta>0$: $\partial f(B_\delta(x_0))\subset Q$. $\forall\ z\in S$，因 $Qz\subset\mathrm{R}^n$ 凸闭，故有 $u\in Qz$: $|u|=d(0,Qz)\geqslant\varepsilon$，于是 $\langle Qz-u,u\rangle\geqslant0$（参考 5.11.1），从而 $w=u/|u|\in S$ 满足 $\langle w,Qz\rangle\geqslant|u|\geqslant\varepsilon$，这得出(12).

令 $W=B_\delta(x_0)$，今证

$$|f(x)-f(y)|\geqslant\varepsilon|x-y|\ (x,y\in W). \tag{13}$$

取定 $x,y\in W$，$x\neq y$，令 $\lambda=|x-y|$，$z=(y-x)/\lambda$. 以 $P$ 记过 $x$ 且正交于 $z$ 的超平面，则对几乎所有 $x'\in P$，$(x'+Rz)\cap W\cap N_f$ 是 1 维零测集. 取如上的 $x'$ 使其充分接近于 $x$，则 $x'+[0,\lambda]z\subset W$，

$$\begin{aligned}|\Delta f(x',\lambda z)|&=\left|\int_0^\lambda f'(x'+tz)z\mathrm{d}t\right|\\&\geqslant\left\langle w,\int_0^\lambda f'(x'+tz)z\mathrm{d}t\right\rangle\\&=\int_0^\lambda\langle w,f'(x'+tz)z\rangle\mathrm{d}t\geqslant\lambda\varepsilon,\end{aligned}$$

其中 $w$ 依(12). 令 $x'\to x$ 得出(13).

令 $V=B(f(x_0),\varepsilon\delta/2)$，$U=W\cap f^{-1}V$. 由(13)，$f:U\to V$ 为单射. $\forall\ y\in V$，$\exists\ x\in\overline{W}$: $|y-f(x)|=d(y,f\overline{W})$. 由(13)有

$$\begin{aligned}|x-x_0|&\leqslant\varepsilon^{-1}|f(x)-f(x_0)|\\&\leqslant\varepsilon^{-1}[|f(x)-y|+|y-f(x_0)|]\\&\leqslant2\varepsilon^{-1}|y-f(x_0)|<\delta,\end{aligned}$$

可见 $x\in W$. 于是 $0\in\partial|y-f(x)|^2=-2(y-f(x))\circ\partial f(x)$(4.8.3). (12)推出 $\partial f(x)\subset\mathrm{GL}(\mathrm{R}^n)$，故 $y=f(x)$，从而 $x\in U$，可见 $f:U\to V$ 为双射. 由(13)有 $\mathrm{Lip}(f^{-1}|V)\leqslant\varepsilon^{-1}$，因此 $f:U\to V$ 为 Lip 同

胚.

如同光滑分析中一样,通过考虑一辅助函数,从反函数定理立即过渡到隐函数定理:

**4.8.6 推论**(隐函数定理) 设 $F:\mathrm{R}^n\times\mathrm{R}^m\to\mathrm{R}^m$ 在点 $(x_0,y_0)$ 邻近为 Lip, $F(x_0,y_0)=0$;对任何 $(A,B)\in\partial F(x_0,y_0)$, $B\in\mathrm{R}^{m\times m}$,有 $\det B\neq 0$,则在 $(x_0,y_0)$ 邻近可从方程 $F(x,y)=0$ 解出 Lip 函数 $y=y(x)$,使得 $y(x_0)=y_0$.

**证** 定义

$$G:\mathrm{R}^n\times\mathrm{R}^m\to\mathrm{R}^n\times\mathrm{R}^m,(x,y)\mapsto(x,F(x,y)),$$

则 $G$ 在 $(x_0,y_0)$ 邻近为 Lip, $G(x_0,y_0)=(x_0,0)$. 对 $F$ 的可微点 $(x,y)$ 有

$$G'(x,y)=\begin{pmatrix} I & 0 \\ F_x(x,y) & F_y(x,y)\end{pmatrix}. \tag{14}$$

结合(1)(14)及对 $\partial F(x_0,y_0)$ 的假定得出 $\partial G(x_0,y_0)\subset\mathrm{GL}(\mathrm{R}^{m+n})$. 由 4.8.5, $G(x,y)$ 在 $(x_0,y_0)$ 邻近有局部 Lip 逆,于是所要结论如同在经典分析中一样得出. □

**4.8.7 例** 设 $f(x,y)=(|x|+y,2x+|y|)$, $x,y\in\mathrm{R}$. 当 $xy\neq 0$ 时有

$$f'(x,y)=\begin{pmatrix}\mathrm{sgn}x & 1\\ 2 & \mathrm{sgn}y\end{pmatrix},$$

于是由(1)得出

$$\partial f(0,0)=\left\{\begin{pmatrix}s & 1\\ 2 & t\end{pmatrix}:s,t\in[-1,1]\right\}\subset\mathrm{GL}(\mathrm{R}^2).$$

由 4.8.5, $f$ 在(0,0)处为局部 Lip 同胚.

# 第五章 非线性最优化

最优化问题的一般形式是：

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D. \tag{P}$$

其中 $f: X \to \overline{\mathrm{R}}$ 是"目标函数"，$g: X \to Y$ 与 $h: X \to Z$ 是"约束函数"，$D \subset X$ 是"约束集"，$Y$ 中已由闭凸锥 $Y_+$ 导入序 $\leqslant$，$Y_- = -Y_+$ 称 $M \triangleq g^{-1}(Y_-) \cap h^{-1}(0) \cap D$ 为(P)的可行集，称每个 $x \in M$ 为可行点，称 $f|M$ 的[局部]极小点为问题(P)的[局部]最优解，称 $\alpha \triangleq \inf\limits_M f(x)$ 为(P)的最优值. 当 $Y = \mathrm{R}^m$ 时约定 $Y_+ = \mathrm{R}^m_+$，$g = (g_i)$；当 $Z = \mathrm{R}^p$ 时令 $h = (h_j)$.

本章中，通常以 $\bar{x}$ 记某个被考察的可行点，且令 $A = g'(\bar{x})$，$B = h'(\bar{x})$，只要它们存在.

问题($P$)对应一个"Lagrange 函数"：

$$L(x, \rho, \lambda, \mu) = (\rho f + \lambda g + \mu h)(x), (x, \rho, \lambda, \mu) \in D \times \mathrm{R}_+ \times Y^*_+ \times Z^*.$$

当 $\rho = 1$ 时在 $L$ 中删去 $\rho$；当 $Y = \{0\}$ 或 $Z = \{0\}$ 时在 $L$ 中删去 $\lambda$ 或 $\mu$. 本章的主要课题是：

(A)通过 Lagrange 函数表示问题(P)之最优解存在的必要或充分条件，这方面的结果是经典 Lagrange 乘子理论的优美推广.

(B)建立问题(P)与一个与之关联的"极大问题"之间的对偶关系，这一方向构成内容丰富的对偶理论.

## §1 Dubovickii-Miljutin 定理

上面提出的最优化问题(P)可进一步抽象为：

$$\min f(x), x \in M_i, 1 \leqslant i \leqslant n, \tag{1}$$

其中 $f:X\to\overline{\mathrm{R}}$，$M_i\subset X$. 令 $M=\bigcap M_i$，设 $\bar{x}\in M\bigcap D_f$. 称

$$F_0\triangleq\{z:\text{当 } 0<t\downarrow 0,w\to z \text{ 时 } \Delta f(\bar{x},tw)<0\}\qquad(2)$$

为 $f$ 在 $\bar{x}$ 的“降锥”. 显然 $F_0$ 为开锥；若 $f'(\bar{x})\neq 0$，则

$$F_0=\{z:\langle f'(\bar{x}),z\rangle<0\},F_0^*=\mathrm{R}_-f'(\bar{x}).\qquad(3)$$

**5.1.1 引理** 设 $k_j(0\leqslant j\leqslant n)$ 是非空凸锥，$j<n$ 时 $K_j$ 开，则 $\bigcap K_j=\varnothing\Leftrightarrow$ 存在不全为零的 $u_j\in K_j^*:\sum u_j=0$.

**证** 若 $\bigcap K_j=\varnothing$，则 $\exists\ m<n:K\triangleq\bigcap_0^m K_j\neq\varnothing,K\bigcap K_{m+1}=\varnothing$. 由分离定理，有 $u\in X^*:u(K_{m+1})<u(K)$，这推出 $u\in K^*=\sum_0^m K_j^*$ (2.2.5)，$u_{m+1}\triangleq -u\in K_{m+1}^*$. 设 $u=\sum_0^m u_j,u_j\in K_j^*$；置 $u_{m+2}=\cdots=u_n=0$，则 $\sum u_j=0,u_j\in K_j^*\ (0\leqslant j\leqslant n)$ 不全为零.

反之，若有不全为零的 $u_j\in K_j^*$；$\sum u_j=0$，则 $\exists\ i<n:u_i\neq 0$；这推出 $\bigcap K_j=\varnothing$，否则对 $x\in\bigcap K_j$ 有 $0=\sum u_j(x)\geqslant u_i(x)>0$！ □

**5.1.2 引理** 若 $\bar{x}$ 是(1)的局部最优解，则

$$F_0\bigcap K_M(\bar{x})=F_0\bigcap\bigcap_{j<n}F(M_j,\bar{x})\bigcap K(M_n,\bar{x})=\varnothing,\qquad(4)$$

记号依 §4.6.

**证** 若 $z\in K_M(\bar{x})$，则 $\exists\ t_k\downarrow 0,z_k\to z:\bar{x}+t_kz_k\in M,f(\bar{x}+t_kz_k)\geqslant f(\bar{x})$，可见 $z\,\overline{\in}\,F_0$. 这结合 4.6.2 得所要证. □

**5.1.3 定理** (Dubovickii-Miljutin, 1965) 设 $K_0=F_0$，$K_j=F(M_j,\bar{x})(1\leqslant j<n)$，$K_n=K(M_n,\bar{x})$. 假定 $K_j(0\leqslant j\leqslant n)$ 非空凸. (i)若 $\bar{x}$ 是(1)的局部最优解，则存在不全为零的 $u_j\in K_j^*\ (0\leqslant j\leqslant n)$；$\sum u_j=0$. (ii)若 $M_i$ 皆凸且满足

$$N\triangleq M_1^\circ\bigcap\cdots\bigcap M_{n-1}^\circ\bigcap M_n\neq\varnothing,\qquad(5)$$

$f$ 严格拟凸且 usc，如(i)所述的 $u_j$ 存在，则 $\bar{x}$ 是(1)的最优解.

**证** (i)直接由 5.1.1 与 5.1.2 得出.

(ii)若 $\bar{x}$ 非最优解，则 $\exists\ x_1\in M:f(x_1)<f(\bar{x})$. 取 $y\in N$，令 $y_\tau=\tau'x_1+\tau y=\bar{x}+z_\tau(\tau'=1-\tau\in(0,1))$，则易验知 $\forall\ t\in(0,1):\bar{x}+$

$tz_\tau=t'\bar{x}+ty_\tau\in N$，这推出 $z_\tau\in\bigcap_1^n K_j$. 取 $\tau$ 充分小，使 $y_\tau$ 充分接近 $x_1$. 由 $f$ usc 当 $w\to z_\tau$ 时 $f(\bar{x}+w)<f(\bar{x})$；而由 $f$ 严格拟凸得出 $\forall\ t\in(0,1)$：$f(\bar{x}+tw)<f(\bar{x})$，因此 $z_\tau\in F_0$，从而 $\bigcap_0^n K_j\neq\varnothing$，与 5.1.1 矛盾. □

5.1.3 是 60 年代寻求统一的最优性条件的种种努力中最典型的结果，它可看作经典 Lagrange 乘子定理的抽象形式：等式 $\sum u_j=0$是抽象的"Euler-Lagrange 方程"，而 $u_j$ 就是抽象的 Lagrange 乘子.

5.1.3 的有效应用有赖于方程 $\sum u_j=0$的具体化，而这依赖于 $K_i^*$ 的具体构造；仅当 $M_i$ 足够限定时，才可望得出 $K_j^*$ 的明显表示. 今考虑问题

$$\min f(x),g(x)\leqslant 0,h(x)=0,x\in D. \tag{6}$$

设 $G=g^{-1}(Y_-)$，$H=h^{-1}(0)$，$\bar{x}\in M\triangleq D\bigcap G\bigcap H$，$D$ 凸且 $D^\circ\neq\varnothing$，$f$，$g$，$h$ 在$\bar{x}$可微. 回忆已约定 $Y_-=-Y_+$，$A=g'(\bar{x})$，$B=h'(\bar{x})$. 其次约定 $K_g=Y^\circ_-+\mathrm{R}g(\bar{x})$，$Q=Y_+^*\bigcap\{g(\bar{x})\}^\perp$；注意 $K_g^*=-Q$.

**5.1.4 引理** $A^{-1}K_g\subset F_G(\bar{x})$；若 $R(A)=Y$，$Y^\circ_+\neq\varnothing$，则 $F_G(\bar{x})=A^{-1}K_g$，$F_G^*(\bar{x})=-A^*Q$.

**证** 若 $z\in A^{-1}K_g$，则$\exists\ y\in Y$，$r\in\mathrm{R}$：$Az=y+rg(\bar{x})$，$y\ll 0$. 于是当 $t\downarrow 0$，$w\to z$ 时

$$\begin{aligned}g(\bar{x}+tw)&=g(\bar{x})+tAz+o(t)\\&=(1+rt)g(\bar{x})+ty+o(t)\leqslant 0,\end{aligned}$$

可见 $z\in F_G(\bar{x})$. 因此 $A^{-1}K_g\subset F_G(\bar{x})$.

其次设 $Y^\circ_+\neq\varnothing$，$R(A)=Y$. 取定 $z\in F_G(\bar{x})$. 若 $z\,\overline{\in}\,A^{-1}K_g$，则 $Az\,\overline{\in}\,K_g$. 因 $K_g$ 是非空开凸集，故由分离定理有 $0\neq\lambda\in Y^*$：$\langle\lambda,Az\rangle>\langle\lambda,K_g\rangle$，这推出 $\lambda\in-K_g^*=Q$，$\langle\lambda,Az\rangle\geqslant 0$. 取 $X$ 的 0-邻域 $U$，使当 $t\downarrow 0$ 时$\bar{x}+t(z+U)\subset G$. 因 $R(A)=Y$，故 $V=AU$ 是 $Y$ 的 0-邻域. $\forall\ x\in U$，当 $t\downarrow 0$ 时有

$$0\geqslant g(\bar{x}+t(z+x))=g(\bar{x})+tA(z+x)+o(t);$$

$$0 \geqslant \langle \lambda, g(\bar{x} + t(z + x)) \rangle = t\langle \lambda, Az + Ax \rangle + o(t),$$

由此推出$\langle \lambda, Ax \rangle \leqslant -\langle \lambda, Az \rangle \leqslant 0$. 这表明$\langle \lambda, V \rangle \leqslant 0$，从而 $\lambda = 0$，得出矛盾. 因此 $z \in A^{-1}K_g$；结合前段所证得 $F_G(\bar{x}) = A^{-1}K_g$. 由 2.2.6，$F_G^*(\bar{x}) = A^* K_g^* = -A^* Q$. □

**5.1.5 引理** 设 $h$ 在$\bar{x}$邻近为 $C^1$ 映射，$R(B) = Z$，则 $K_H(\bar{x}) = N(B)$，$K_H^*(\bar{x}) = R(B^*)$.

**证** 若 $z \in K_H(\bar{x})$，则$\exists$ $t_k \downarrow 0, z_k \to z: \bar{x} + t_k z_k \in H$，于是

$$0 = h(\bar{x} + t_k z_k) = t_k B z_k + o(t_k),$$

这推出 $Bz = 0$. 可见 $K_H(\bar{x}) \subset N(B)$. 反之，若 $z \in N(B)$，则对任给 $t_k \downarrow 0$，由 Lusternik 定理(如见[34;p. 94])，方程 $h(\bar{x} + y_k + t_k z) = 0$ 有解 $y_k = o(t_k)$，于是$z_k \triangleq t_k^{-1} y_k + z \to z$，$\bar{x} + t_k z_k \in H$，因此 $z \in K_H(\bar{x})$. 故得 $K_H(\bar{x}) = N(B)$；从而 $K_H^*(\bar{x}) = N(B)^{\perp} = R(B^*)$. □

利用 5.1.4 与 5.1.5 可得出 5.1.3 的如下具体化.

**5.1.6 定理** 设 $Y^\circ_+ \neq \varnothing$，$h$ 在$\bar{x}$邻近为 $C^1$ 映射，$R(B)$闭. 若$\bar{x}$是问题(6)的局部最优解，则有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathrm{R}_+ \times Q \times Z^*$，使得

$$\langle L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}), D - \bar{x} \rangle \geqslant 0, \tag{7}$$

其中 $L(\cdot, \bar{\rho}, \bar{\lambda}, \bar{\mu}) = \bar{\rho} f + \bar{\lambda} g + \bar{\mu} h$. 若 $\bar{x} \in D^\circ$，则(7)成为

$$L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) = 0. \tag{8}$$

**证** 若 $f'(\bar{x}) = 0$，则取$\bar{\rho} = 1, \bar{\lambda} = 0, \bar{\mu} = 0$；若$K_1 \triangleq A^{-1}K_g = \varnothing$，则由 Gordan 定理(参看 5.5.11)有 $0 \neq \bar{\lambda} \in -K_g^* = Q: A^* \bar{\lambda} = 0$，取$\bar{\rho} = 0, \bar{\mu} = 0$；若 $R(B) \neq Z$，则取 $0 \neq \bar{\mu} \in N(B^*)$，$\bar{\rho} = 0, \bar{\lambda} = 0$. 以上取法皆使 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathrm{R}_+ \times Q \times Z^*$ 满足(8). 因此下面不妨设 $f'(\bar{x}) \neq 0, K_1 \neq \varnothing, R(B) = Z$. 令 $K_0 = \{z: \langle f'(\bar{x}), z \rangle < 0\}$，$K_2 = F_D(\bar{x})$，$K_3 = K_H(\bar{x})$. 则 $K_0, K_1, K_2$ 为非空开凸锥；$K_0 = F_0$，$K_0^* = \mathrm{R}_- f'(\bar{x})$(见(3))；$K_1^* = -A^* Q$(参照 5.1.4 之证)；$K_2 = \mathrm{cone}(D^\circ - \bar{x})$，$K_2^* = (D - \bar{x})^*$ (4.6.2(i))；$K_3 = N(B)$，$K_3^* = R(B^*)$ (5.1.5). 由 5.1.4 有 $K_1 \subset F_G(\bar{x})$；由 5.1.2 有 $F_0 \cap K_M(\bar{x}) = \varnothing$，

于是(参考 4.6.2)

$$\phi = K_0 \cap F_G(\bar{x}) \cap F_D(\bar{x}) \cap K_H(\bar{x}) = K_0 \cap K_1 \cap K_2 \cap K_3.$$

由 5.1.1,有不全为零的 $u_j \in K_j^*$ $(0 \leqslant j \leqslant 3)$: $\sum u_j = 0$. 设 $u_0 = -\bar{\rho} f'(\bar{x}), \bar{\rho} \in \mathrm{R}_+$; $u_1 = -A^* \bar{\lambda}, \bar{\lambda} \in Q$; $u_3 = -B^* \bar{\mu}, \bar{\mu} \in Z^*$, 则$(\bar{\rho}, \bar{\lambda}, \bar{\mu}) \neq 0$,

$$0 \leqslant \langle u_2, D - \bar{x} \rangle = \langle -u_0 - u_1 - u_3, D - \bar{x} \rangle =$$
$$\langle L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}), D - \bar{x} \rangle. \qquad \square$$

若 $Y = \mathrm{R}^m$, $g = (g_i)$, 则令 $I = \{i : g_i(\bar{x}) = 0\}$, $g_I = (g_i)_{i \in I}$, $A_I = g'_I(\bar{x})$; 今后将保持这些记号. 注意此时 $Q = \{(\lambda_i) \in \mathrm{R}_+^m : i \overline{\in} I \Rightarrow \lambda_i = 0\} \triangleq \mathrm{R}_+^I$, 于是从 5.1.6 得:

**5.1.7 推论** 设 $Y = \mathrm{R}^m$, $h$ 在$\bar{x}$邻近为 $C^1$ 映射, $R(B)$闭. 若$\bar{x}$是问题(6)的局部最优解, 则存在 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathrm{R}_+ \times R_+^I \times Z^*$ 满足(7).

通常称 5.1.7 为"Fritz John 定理".

## §2 Kuhn-Tucker 条件

本节考虑问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0. \tag{1}$$

设 $G = g^{-1}(Y_-)$, $H = h^{-1}(0)$, $\bar{x} \in M = G \cap H$, $L(\cdot, \lambda, \mu) = f + \lambda g + \mu h$. 假定 $f, g$ 在$\bar{x}$可微, $h$ 在$\bar{x}$邻近为 $C^1$ 映射.

**5.2.1 定义** 若存在$(\bar{\lambda}, \bar{\mu}) \in Y_+^* \times Z^*$, 使得

$$L_x(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0, \langle \bar{\lambda}, g(\bar{x}) \rangle = 0, \tag{2}$$

则称$\bar{x}$为问题(1)的 Kuhn-Tucker 点, 称$\bar{\lambda}, \bar{\mu}$为 Kuhn-Tucker 乘子, 称(2)为 Kuhn-Tucker 条件(分别简称为 *K-T* 点、*K-T* 乘子与 *K-T* 条件).

形式上, 5.2.1 并未涉及$\bar{x}$是否为最优解. 从 Kuhn, Tucker 开始的一系列研究致力于寻求加于约束的一定条件——通常称为

“约束品性”——以保证局部最优点是 $K$-$T$ 点.

首先注意到，若依 5.1.6 得出 §1(8)且其中$\bar{\rho}\neq 0$，则$\bar{x}$显然为 $K$-$T$ 点. 据此可证：

**5.2.2 定理** 设 $Y^{\circ}_{+}\neq\varnothing, R(B)=Z$. 若$\bar{x}$是(1)的局部最优点，则以下每个条件推出$\bar{x}$为 $K$-$T$ 点：(i)$(A,B):X\to Y\times Z$ 为满射；(ii)$N(B)\cap A^{-1}(K_g)\neq\varnothing, K_g=Y^{\circ}_{-}+\mathrm{R}g(\bar{x})$；(iii)$g$ 为 * 伪凸(4.1.11)，$h(x)=Bx-b, \exists\ \hat{x}\in H: g(\hat{x})\ll 0$.

**证** 由 5.1.6，有 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in \mathrm{R}_{+}\times Q\times Z^{*}$，使得 §1(8)成立，此处及以下都约定 $Q=Y^{*}_{+}\cap\{g(\bar{x})\}^{\perp}$. 于是只需证$\bar{\rho}\neq 0$. 假定$\bar{\rho}=0$，则$(\bar{\lambda},\bar{\mu})\neq 0, A^{*}\bar{\lambda}+B^{*}\bar{\mu}=0$. 可设$\bar{\lambda}\neq 0$(否则 $0\neq\bar{\mu}\in N(B^{*})$，与 $R(B)=Z$ 矛盾). 若条件(i)满足，则$(A,B)^{*}$为单射，与 $A^{*}\bar{\lambda}+B^{*}\bar{\mu}=0$ 矛盾. 若条件(ii)满足，则$\exists\ z\in N(B): Az\in K_g$，于是

$$0=\langle A^{*}\bar{\lambda}+B^{*}\bar{\mu},z\rangle=\langle\bar{\lambda},Az\rangle<0,$$

得出矛盾. 若条件(iii)满足，令 $z=\hat{x}-\bar{x}$，则 $Bz=0$，于是$\langle\bar{\lambda}g'(\bar{x}),\hat{x}-\bar{x}\rangle=\langle\bar{\lambda},Az\rangle=0$，从而由$\bar{\lambda}g$ 伪凸得出$\langle\bar{\lambda},g(\hat{x})\rangle\geqslant\langle\bar{\lambda},g(\bar{x})\rangle=0$，这与 $g(\hat{x})\ll 0$ 矛盾. □

若 $Y=\mathrm{R}^m$，则对邻近$\bar{x}$ 的 $x$ 有 $g(x)\leqslant 0\Leftrightarrow g_I(x)\leqslant 0$(回忆上节中的记号约定)，因此，对于局部问题可以 $g_I$ 替换(1)中的 $g$. 据此，从 5.2.2 得到：

**5.2.3 推论** 设 $Y=\mathrm{R}^m, R(B)=Z$，$\bar{x}$ 是(1)的局部最优解，则以下每个条件推出$\bar{x}$为 $K$-$T$ 点：(i)$(A_I,B):X\to\mathrm{R}^I\times Z$ 为满射；(ii)$\exists\ z\in N(B): A_I z\ll 0$；(iii)$g_I$ 在$\bar{x}$相对于 $M$ 为 * 伪凸，$h(x)=Bx-b, \exists\ \hat{x}\in M: g_I(\hat{x})\ll 0$.

**注** 若 $Z=\mathrm{R}^p, h=(h_j)$，则 5.2.3 之条件(i)相当于$\{g'_i(\bar{x}); i\in I\}\cup\{h'_j(\bar{x}): 1\leqslant j\leqslant p\}$线性无关.

得出 $K$-$T$ 点的另一途径基于以下事实：$\bar{x}$是问题(1)的 $K$-$T$ 点$\Leftrightarrow\exists\ (\bar{\lambda},\bar{\mu})\in Q\times Z^{*}: -f'(\bar{x})=A^{*}\bar{\lambda}+B^{*}\bar{\mu}\Leftrightarrow -f'(\bar{x})\in A^{*}Q+R(B^{*})$. 因此，若能求得一集 $K\subset X$，使之满足

$$F_0 \cap K = \varnothing, \; -K^* \subset A^*Q + R(B^*). \tag{3}$$

其中 $F_0$ 依 §1(2)，则 $\bar{x}$ 必为 $K$-$T$ 点. 否则，必 $f'(x) \overline{\in} K^*$，于是 $\exists\, z \in K$：$\langle f'(\bar{x}), z\rangle < 0$，从而 $z \in F_0 \cap K$(参考 §1(3))！基于以上讨论，可证

**5.2.4 定理** 设 $A^*Q + R(B^*)$ 弱$^*$闭，$K = \{z \in N(B): \langle Q, Az\rangle \leqslant 0\}$，$F_0 \cap K = \varnothing$，则 $\bar{x}$ 是(1)的 $K$-$T$ 点.

**证** 只需验证(3)中的包含. 设 $u \in X^*$. 若 $u \overline{\in} A^*Q + R(B^*)$，则由分离定理有 $z \in X$：

$$u(z) > \langle A^*Q + R(B^*), z\rangle = \langle Q, Az\rangle + \langle Z^*, Bz\rangle.$$

这推出 $u(z) > 0$，$\langle Q, Az\rangle \leqslant 0$，$Bz = 0$(用 2.2.2)，可见 $z \in K$，$u \overline{\in} -K^*$. □

所宜注意者，5.2.4 并不直接涉及 $\bar{x}$ 是否为最优解. 为有效应用 5.2.4，集 $Q$ 与 $K$ 需有更具体的表示. 考虑 $Y = \mathrm{R}^m$ 这一特殊情况. 如上节所指出的，在这种情况下有 $Q = \mathrm{R}^I_+$. 相应地，5.2.4 中的 $K$ 可表成：

$$K = \{z \in N(B): A_I z \leqslant 0\} = N(B) \cap A_I^{-1}\mathrm{R}^I_-. \tag{4}$$

利用 $Q$ 与 $K$ 的以上表示，易从 5.2.4 得出：

**5.2.5 推论** 设 $Y = \mathrm{R}^m$，$\bar{x}$ 是(1)的局部最优解，$A_I^*\mathrm{R}^I_+ + R(B^*)$ 弱$^*$闭，则以下每个条件推出 $\bar{x}$ 为 $K$-$T$ 点：(i) $N(B) \cap A_I^{-1}\mathrm{R}^I_- \subset K_M(\bar{x})$；(ii) $g_i\,(i \in I)$ 伪凸，$h(x) = Bx - b$，$\exists\, \hat{x} \in H$：$g_I(\hat{x}) \ll 0$.

**证** 设 $K$ 依(4)，只需指明 $F_0 \cap K = \varnothing$. 当条件(i)满足时这由 5.1.2 得出. 以下设条件(ii)满足. 不妨设 $I = \{1, 2, \cdots, m\}$，即 $g(\bar{x}) = 0$. 令 $K_0 = \{z \in N(B): Az \ll 0\}$，则由 5.1.4 有 $K_0 \subset N(B) \cap F_G(\bar{x})$. 若 $z \in K_0$，则 $Bz = 0$，当 $t \downarrow 0$ 时 $\bar{x} + tz \in G$，$h(\bar{x} + tz) = B\bar{x} - b = 0$，因此 $\bar{x} + tz \in M$，从而 $f(\bar{x} + tz) \geqslant f(\bar{x})$，于是 $z \overline{\in} F_0$. 这表明 $F_0 \cap K_0 = \varnothing$. 下面只要证 $K = \overline{K}_0$. 令 $z_0 = \hat{x} - \bar{x}$，则 $Bz_0 = 0$；用 $g_i$ 伪凸及 $g(\hat{x}) \ll 0$ 易得 $Az_0 \ll 0$，因此 $z_0 \in K_0$. $\forall\, z \in K$，$t \in (0,$

1)，显然$z_t \triangleq tz_0+t'z\in K_0$，而$z_t\to z(t\downarrow 0)$，这表明$K=\overline{K}_0$. □

**注** 易验知恒有$K_M(\bar{x})\subset K$，故5.2.5之条件(i)相当于$K_M(\bar{x})=K$. 5.2.5中"弱*闭条件"之验证颇成问题. 不过，当$\dim X<\infty$时无需验证此条件.

对于问题(1)的以下特殊情形：

$$\min f(x),g(x)\leqslant 0 \tag{5}$$

与

$$\min f(x),h(x)=0, \tag{6}$$

相应于5.2.2～5.2.5的结果可表述得简单些. 首先注意，$\bar{x}\in G$是(5)的$K-T$点$\Leftrightarrow f'(\bar{x})\in -A^*Q$；$\bar{x}\in H$是(6)的$K$-$T$点$\Leftrightarrow f'(\bar{x})\in R(B^*)$. 在前述结果中分别取$Z=\{0\}$与$Y=\{0\}$得出：

**5.2.6 推论** 设$\bar{x}$是(5)的局部最优解，则以下每个条件推出$\bar{x}$为$K$-$T$点：(i)$Y^{\circ}_+\neq\varnothing$，$R(A)=Y$；(ii)$Y^{\circ}_+\neq\varnothing$，$g$为*伪凸，$\exists\,\hat{x}\in X$：$g(\hat{x})\ll 0$，(iii)$Y=\mathrm{R}^m$，$R(A_I)=\mathrm{R}^I$；(iv)$Y=\mathrm{R}^m$，$g_i(i\in I)$伪凸，$\exists\,\hat{x}\in X$：$g(\hat{x})\ll 0$.

**5.2.7 推论** 设$\bar{x}$是(6)的局部最优解，则以下每个条件推出$\bar{x}$为$K$-$T$点：(i)$R(B)=Z$；(ii)$R(B)$闭，$F_0\cap N(B)=\varnothing$；(iii)$R(B)$闭，$N(B)\subset K_H(\bar{x})$.

**注** 由闭值域定理[198]，$R(B)$闭$\Rightarrow R(B^*)$弱*闭.

下面转而讨论$K$-$T$条件对于最优解的充分性. 我们希望当$f,g,h$满足适当"凸性"条件时(1)的$K$-$T$点是最优解. 现代研究已将所需"凸性"条件降低到很宽的程度，这依赖于以下概念：

**5.2.8 定义** 设$\eta:D\subset X\to X$. 若$\forall\, x\in D\setminus\{\bar{x}\}$：$\langle f'(\bar{x}),\eta(x)\rangle\leqslant[<]f(x)-f(\bar{x})$，则说$f$在$\bar{x}$相对于$D$为$\eta$-凸[$\eta$-严格凸]；若$\forall\, x\in D$：$\langle f'(\bar{x}),\eta(x)\rangle>0\Rightarrow f(x)>f(\bar{x})$，则说$f$在$\bar{x}$相对于$D$为$\eta$-拟凸；若$\forall\, x\in D\setminus\{\bar{x}\}$：$\langle f'(\bar{x}),\eta(x)\rangle\geqslant 0\Rightarrow f(x)\geqslant f(\bar{x})$[$f(x)>f(\bar{x})$]，则说$f$在$\bar{x}$相对于$D$为$\eta$-伪凸[$\eta$-严格伪凸].

若取$\eta(x)=x-\bar{x}$，则$\eta$-凸重合于凸(4.1.5)；$\eta$-拟凸等仿此. 显然$\eta$-严格凸$\Rightarrow\eta$-凸$\Rightarrow\eta$-伪凸$\Leftarrow\eta$-严格伪凸$\Leftarrow\eta$-严格凸；$\eta$-凸$\Rightarrow\eta$-

拟凸. 若 $f:\mathrm{R}\to\mathrm{R}$, $f'(\bar{x})\neq 0$, 令 $\eta(x)\doteq[f(x)-f(\bar{x})]/f'(\bar{x})$, 则 $f$ 在 $\bar{x}$ 为 $\eta$-凸. 由此例足见"$\eta$-凸"的宽泛性.

**5.2.9 定理** 设 $\bar{x}$ 是(1)的 $K$-$T$ 点, $\bar{\lambda}$, $\bar{\mu}$ 是 $K$-$T$ 乘子. (i)若 $f$ 与 $\bar{\lambda}g+\bar{\mu}h$ 在 $\bar{x}$ 相对于 $M$ 分别为 $\eta$-伪凸与 $\eta$-拟凸, 则 $\bar{x}$ 是(1)的最优解. (ii)若 $f$ 与 $\bar{\lambda}g+\bar{\mu}h$ 在 $\bar{x}$ 相对于 $M$ 分别为 $\eta$-拟凸与 $\eta$-严格伪凸(或分别为 $\eta$-严格伪凸与 $\eta$-拟凸), 则 $\bar{x}$ 是(1)的严格最优解(这意味着 $\forall\ x\in M\setminus\{\bar{x}\}$: $f(x)>f(\bar{x})$).

**证** (i)若 $\bar{x}$ 非最优解, 则 $\exists\ x\in M$: $f(x)<f(\bar{x})$. 由 $f\eta$一伪凸有 $\langle f'(\bar{x}),\eta(x)\rangle<0$, 从而由(2)得 $\langle\bar{\lambda}\circ A+\bar{\mu}\circ B,\eta(x)\rangle>0$, 于是由 $\bar{\lambda}g+\bar{\mu}h$ 为 $\eta$-拟凸得出

$$0\geqslant(\bar{\lambda}g+\bar{\mu}h)(x)>(\bar{\lambda}g+\bar{\mu}h)(\bar{x})=0,$$

导致矛盾.

(ii)的证明是类似的. □

5.2.9 中, 函数 $\eta$ 的具体形式不起任何作用, 重要的只是其存在性(自然要求在同一问题中 $\eta$ 保持固定), 这种存在性甚至可在不提及 $\eta$ 的形式下表达出来. 例如, $f$, $\bar{\lambda}g$ 与 $\bar{\mu}h$ 在 $\bar{x}$ 相对于 $M$ 为 $\eta$-凸 $\Leftrightarrow\forall\ x\in M$, $\exists\ \eta(x)\in X$: $F'(x)\eta(x)\leqslant F(x)-F(\bar{x})$, 此处 $F=(f,\bar{\lambda}g,\bar{\mu}h)$. 由 5.5.13, 后者等价于 $\forall\ x\in M$, 问题

$$\exists\ \alpha\in\mathrm{R}^3_+: F'(\bar{x})^*\alpha=0,\ \langle\alpha,Fx-F\bar{x}\rangle=-1 \tag{7}$$

无解. 令 $\alpha=(\alpha_i)$, 利用 $-f'(\bar{x})=A^*\bar{\lambda}+B^*\bar{\mu}$ 可将(7)写成:

$$\begin{cases}\alpha_1[f(x)-f(\bar{x})]+\alpha_2\bar{\lambda}g(x)+\alpha_3\bar{\mu}h(x)=-1;\\(\alpha_2-\alpha_1)A^*\bar{\lambda}=(\alpha_1-\alpha_3)B^*\bar{\mu},\alpha_1,\alpha_2,\alpha_3\geqslant 0.\end{cases} \tag{8}$$

**5.2.10 例** 令 $X=\mathrm{R}^2$, $Y=\mathrm{R}^6$, $Z=\{0\}$. 在(1)中取 $f(x)=x_1-\sin x_2$, $g=(g_i)$, $g_1(x)=\sin x_1-4\sin x_2$, $g_2(x)=2\sin x_1+7\sin x_2+x_1-6$, $g_3(x)=2(x_1+x_2)-3$, $g_4(x)=4|x|^2-9$, $g_5(x)=-\sin x_1$, $g_6(x)=-\sin x_2$. 直接验知 $\bar{x}=(0,\arcsin(6/7))$ 是 $K$-$T$ 点, $\bar{\lambda}=(0,1/7,0,0,10/7,0)$. 因 $g'(\bar{x})^*\bar{\lambda}=(g'_2(\bar{x})+10g'_5(\bar{x}))/7\neq 0$, 故从(8)的第二个方程得 $\alpha_1=\alpha_2$. 由此易见(8)对 $x\in M\triangleq$

$g^{-1}(\mathrm{R}_-^c)$无解，从而 $f$ 与 $\bar{\lambda}g$ 在 $\bar{x}$ 相对于 $M$ 为 $\eta$-凸（对某个 $\eta$），于是由 5.2.9 知 $\bar{x}$ 是最优解. 这一结论亦可加以直接验证.

## §3　2 阶 条 件

上节结果表明，关于问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0 \tag{1}$$

的最优性条件略同于 Larange 函数 $L(\cdot, \lambda, \mu)$的无约束极小条件. 本节在 2 阶导数的范围内确立类似结论. 以下设 $G = g^{-1}(Y_-)$, $H = h^{-1}(0)$, $\bar{x} \in M = G \cap H$, $f, g, h$ 在 $\bar{x}$ 邻近为 $C^2$ 函数, $Q = Y_+^* \cap \{g(\bar{x})\}^\perp$.

**5.3.1 定理**　设 $\bar{x}$ 是(1)的 $K$-$T$ 点, $\bar{\lambda}, \bar{\mu}$ 为 $K$-$T$ 乘子. (i)必要条件：若 $\bar{x}$ 是(1)的局部最优解, $N = \{x \in M : \langle \bar{\lambda}, g(x) \rangle = 0\}$, 则

$$L_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu})z^2 \geqslant 0 (\forall z \in K_N(\bar{x})). \tag{2}$$

(ii)充分条件：若 $X = \mathrm{R}^n$, $K = \{z \in N(B) : \langle Q, Az \rangle \leqslant 0 = \langle \bar{\lambda}, Az \rangle\}$,

$$L_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu})z^2 > 0 (\forall z \in K \cap S^{n-1}), \tag{3}$$

则 $\bar{x}$ 是(1)的严格局部最优解. (iii)充分条件：若 $X = \mathrm{R}^n$, 存在 $\delta > 0$, 使得

$$L_{xx}(x, \bar{\lambda}, \bar{\mu})z^2 \geqslant 0 (x \in B_\delta(\bar{x}), d(z, K \cap S^{n-1}) < \delta), \tag{4}$$

其中 $K$ 依(ii), 则 $\bar{x}$ 是(1)的局部最优解.

**证**　(i)若 $z \in K_N(\bar{x})$, 则有 $t_k \downarrow 0$, $z_k \to z$, 使 $x_k \triangleq \bar{x} + t_k z_k \in N$ (§4.6(4)), 于是 $\langle \bar{\lambda}, g(x_k) \rangle = \langle \bar{\mu}, h(x_k) \rangle = 0$. 将 $L(\cdot, \bar{\lambda}, \bar{\mu})$ 作 2 阶 Taylor 展开并用 $K$-$T$ 条件得：

$$0 \leqslant f(x_k) - f(\bar{x}) = L(x_k, \bar{\lambda}, \bar{\mu}) - L(\bar{x}, \bar{\lambda}, \bar{\mu})$$

$$= \frac{1}{2} t_k^2 L_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu}) z_k^2 + o(t_k^2),$$

这推出 $L_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu})z^2 \geqslant 0$. 可见条件(2)满足.

(ii)若 $\bar{x}$ 非严格局部最优解，则∃ $x_k \in M \setminus \{\bar{x}\}$ : $x_k \to \bar{x}$, $f(x_k) \leqslant f(\bar{x})$. 令 $x_k = \bar{x} + t_k z_k$, $t_k = |x_k - \bar{x}| \to 0$, $|z_k| = 1$, 不妨设 $z_k \to z \in$

$S^{n-1}$. 分别由

$$0 \geqslant f(x_k) - f(\bar{x}) = t_k\langle f'(\bar{x}), z_k\rangle + o(t_k),$$

$$0 \geqslant \langle Q, g(x_k) - g(\bar{x})\rangle = \langle Q, t_k A z_k + o(t_k)\rangle$$

与
$$0 = h(x_k) - h(\bar{x}) = t_k B z_k + o(t_k)$$

得$\langle f'(\bar{x}), z\rangle \leqslant 0$, $\langle Q, Az\rangle \leqslant 0$ 与 $Bz = 0$. 这结合

$$0 = \langle L_x(\bar{x}, \bar{\lambda}, \bar{\mu}), z\rangle = \langle f'(\bar{x}), z\rangle + \langle \bar{\lambda}, Az\rangle$$

得$\langle f'(\bar{x}), z\rangle = \langle \bar{\lambda}, Az\rangle = 0$，因此 $z \in K \cap S^{n-1}$. 但由

$$\begin{aligned} 0 &\geqslant f(x_k) - f(\bar{x}) = L(x_k, \bar{\lambda}, \bar{\mu}) - L(\bar{x}, \bar{\lambda}, \bar{\mu}) - \langle \bar{\lambda}, g(x_k)\rangle \\ &\geqslant L(x_k, \bar{\lambda}, \bar{\mu}) - L(\bar{x}, \bar{\lambda}, \bar{\mu}) \\ &= \frac{1}{2} t_k^2 L_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu}) z_k^2 + o(t_k^2) \end{aligned}$$

推出 $L_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu}) z^2 \leqslant 0$，与条件(3)矛盾.

(iii)若 $\bar{x}$ 非局部最优解，则$\exists$ $x_k \in M \backslash \{\bar{x}\}$: $x_k \to \bar{x}$, $f(x_k) < f(\bar{x})$. 由(ii)之证，可设 $x_k = \bar{x} + t_k z_k$, $t_k \downarrow 0$, $z_k \to z \in K \cap S^{n-1}$. 另一方面有

$$\begin{aligned} 0 &> f(x_k) - f(\bar{x}) \geqslant L(x_k, \bar{\lambda}, \bar{\mu}) - L(\bar{x}, \bar{\lambda}, \bar{\mu}) \\ &= \frac{1}{2} t_k^2 L_{xx}(y_k, \bar{\lambda}, \bar{\mu}) z_k^2, \end{aligned} \tag{5}$$

其中 $y_k \in [\bar{x}, x_k]$. 当 $k$ 充分大时 $y_k \in B_\delta(\bar{x})$, $d(z_k, K \cap S^{n-1}) \leqslant |z_k - z| < \delta$，故由条件(4)有 $L_{xx}(y_k, \bar{\lambda}, \bar{\mu}) z_k^2 \geqslant 0$，这与(5)矛盾.

**注** 若 $Y = \mathbb{R}^m$，则 5.3.1 中的 $N, K$ 可表成：

$$N = \{x \in M : \bar{\lambda}_i g_i(x) = 0 (\forall\, i \in I)\};$$

$$K = \{z \in N(B) : A_I z \leqslant 0, \langle \bar{\lambda}_i g'_i(\bar{x}), z\rangle = 0 (\forall\, i \in I)\},$$

有关记号参看§1.

对于纯等式约束问题

$$\min f(x), h(x) = 0, \tag{6}$$

令 $L(\cdot, \mu) = f + \mu h$，在 5.3.1 中取 $Y = \{0\}$ 得出

**5.3.2 推论** 设 $\bar{\mu} \in Z^*$, $L_x(\bar{x}, \bar{\mu}) = 0$. (i)若 $\bar{x}$ 是(6)的局部最优解，则

$$L_{xx}(\bar{x},\bar{\mu})z^2 \geqslant 0(\forall\ z \in K_H(\bar{x})). \tag{7}$$

(ii)若 $X=\mathbf{R}^n$,

$$L_{xx}(\bar{x},\bar{\mu})z^2 > 0(\forall\ z \in N(B) \cap S^{n-1}), \tag{8}$$

则 $\bar{x}$ 是(6)的严格局部最优解.(iii)若 $X=\mathbf{R}^n$,存在 $\delta>0$,使得

$$L_{xx}(x,\bar{\mu})z^2 \geqslant 0(x \in B_\delta(\bar{x}), d(z,N(B) \cap S^{n-1}) < \delta), \tag{9}$$

则 $\bar{x}$ 是(6)的局部最优解.

**注** 若 $R(B)=Z$,$\bar{x}$ 是(6)的局部最优解,则必有 $\bar{\mu}\in Z^*$ 使 $L_x(\bar{x},\bar{\mu})=0$. 且 $K_H(\bar{x})=N(B)$(5.2.7,5.1.5).

原则上,本节的方法可进一步向高阶拓广,但必然出现技术性的困难.下面仅建立一个较特殊的结果,为此需要以下概念:

**5.3.3 定义** 若 $F:\Omega\subset X\to Y$ 是 $C^r(r\geqslant 1)$映射,$x_0\in\Omega$,$F'(x_0)$是核裂的(即有分解 $X=X_1\oplus N(F'(x_0))$),且 $R(F'(x_0))=Y$,则说 $F$ 在 $x_0$ 为浸满(或 $C^r$ 浸满).

**注** $C^r(r\geqslant 1)$映射 $F=(F_i):\Omega\subset\mathbf{R}^n\to\mathbf{R}^m$ 在 $x\in\Omega$ 为浸满$\Leftrightarrow$ $\{F'_i(x):1\leqslant i\leqslant m\}$线性无关.

**5.3.4 引理** 若 $h$ 在 $\bar{x}\in H$ 为 $C^r(r\geqslant 1)$浸满,则存在从 $N(B)$ 的一 0-邻域 $V$ 到 $\bar{x}$ 在 $H$ 中的某邻域的同胚 $\varphi$,使当 $z\in V$,$z\to 0$ 时 $\varphi(z)=\bar{x}+z+o(z)$,且 $\varphi:V\to X$ 为 $C^r$ 函数.

**证** 设 $X=X_1\oplus X_2$,$X_1=N(B)$. 对方程

$$F(x_1,x_2) \triangleq h(\bar{x}+x_1+x_2) = 0(x_i \in X_i, i=1,2)$$

用隐函数定理,可在 $0\in X$ 邻近局部地解出 $C^r$ 函数 $x_2=\psi(x_1)$,使 $\psi(0)=0$. 直接算出

$$\psi'(0) = -(F'_{x_2}(0,0))^{-1}F'_{x_1}(0,0) = 0. \tag{10}$$

令 $\varphi(z)=\bar{x}+z+\psi(z)$,则 $\varphi$ 是 $X_1$ 的某 0-邻域 $V$ 上的 $C^r$ 函数,$\varphi(0)=\bar{x}$,$\varphi(z)\in H$;(10)表明 $\varphi(z)=\bar{x}+z+0(z)(z\to 0)$. 设 $P:X\to X_1$ 是投影,则 $P(\varphi(z)-\bar{x})=z$,由此看出 $\varphi$是所求的同胚. □

利用 5.3.4 可得如下高阶充分条件:

**5.3.5 定理** 设 $\bar{x}\in H$,$f$、$h$ 在 $\bar{x}$ 邻近为 $C^r$ 函数,$r\geqslant 2$ 为偶

数;$h$ 在 $\bar{x}$ 为浸满. 若存在 $\bar{\mu}\in Z^*$,$c>0$,使得 $L_x^{(s)}(\bar{x},\bar{\mu})=0(1\leqslant s<r)$,

$$L_x^{(r)}(\bar{x},\bar{\mu})z^r\geqslant c|z|^r(\forall\ z\in N(B)),\tag{11}$$

则 $\bar{x}$ 是(6)的严格局部最优解.

**证** 若 $\bar{x}$ 非严格局部最优解,则有 $x_k\in H\setminus\{\bar{x}\}$:$x_k\to\bar{x}$,$f(x_k)\leqslant f(\bar{x})$. 设 $\varphi$ 如 5.3.4,令 $z_k=\varphi^{-1}(x_k)$,则 $0\neq z_k\in N(B)$,$z_k\to 0$,$\varphi(z_k)=\bar{x}+z_k+o(z_k)$. 利用 $L(\cdot,\bar{\mu})$ 在 $\bar{x}$ 的 $r$ 阶 Taylor 展开与(11)得出,当 $k\to\infty$ 时有

$$\begin{aligned}0&\geqslant f(x_k)-f(\bar{x})=L(x_k,\bar{\mu})-L(\bar{x},\bar{\mu})\\&=\frac{1}{r!}L_x^{(r)}(\bar{x},\bar{\mu})(x_k-\bar{x})^r+o(|x_k-\bar{x}|^r)\\&=\frac{1}{r!}L_x^{(r)}(\bar{x},\bar{\mu})z_k^r+o(|z_k|^r)\geqslant\frac{c}{r!}|z_k|^r+o(|z_k|^r)>0,\end{aligned}$$

得出矛盾. □

## §4 非光滑最优性条件

本节对于问题

$$\min f(x),g(x)\leqslant 0,h(x)=0,x\in D\tag{1}$$

给出 Frite John 定理(5.1.7)的两个非光滑类似. 设 $Y=\mathrm{R}^m$,$G=g^{-1}(\mathrm{R}_-^m)$,$H=h^{-1}(0)$,$\bar{x}\in M=D\cap G\cap H$,$L(\cdot,\rho,\lambda,\mu)=\rho f+\lambda g+\mu h$;假定 $f,g,h$ 在 $\bar{x}$ 邻近为 Lip.

第一个结果基于 5.1.1.

**5.4.1 定理** 设 $D$ 凸,$D^\circ\neq\varnothing$;$h$ 在 $\bar{x}$ 邻近为 $C^1$ 函数,$R(B)$ 闭. 若 $\bar{x}$ 是(1)的局部最优解,则有 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in\mathrm{R}_+\times\mathrm{R}_+^I\times Z^*$(记号依 §1),$\bar{\lambda}=(\bar{\lambda}_i)$,$\bar{u}_0\in\partial f(\bar{x})$,$\bar{u}_i\in\partial g_i(\bar{x})$,使得

$$\langle\bar{\rho}u_0+\sum\bar{\lambda}_i\bar{u}_i+B^*\bar{u},D-\bar{x}\rangle\geqslant 0.\tag{2}$$

若 $\bar{x}\in D^\circ$,$R(B)=Z$,则当以下条件之一满足时 $\bar{\rho}\neq 0$:(i)$\exists\ z\in N(B)$,$\forall\ i\in I$:$g^\circ_i(\bar{x},z)<0$;(ii)$g_i(i\in I)$在 $\bar{x}$ 相对于 $M$ 广义伪凸

(参看 § 4.3), $h(x)=Bx-b$, $\exists\ \hat{x}\in M: g_I(\hat{x})\ll 0$.

**证** 不妨设 $g(\bar{x})=0$(否则以 $g_I$ 替换 $g$),且可设 $0\overline{\in}\partial f(\bar{x})\cup\partial g_i(\bar{x})(1\leqslant i\leqslant m)$, $R(B)=Z$. 令 $K_0=\{z: f^\circ(\bar{x},z)<0\}$, $K_i=\{z: g^\circ_i(\bar{x},z)<0\}(1\leqslant i\leqslant m)$, $K_{m+1}=F_D(\bar{x})$, $K_n=N(B)=K_H(\bar{x})$, $n=m+2$. 设 $F_0$ 依 § 1(2), $z\in X\backslash F_0$, 则 $\exists\ t_k\downarrow 0, z_k\to z, x_k=\bar{x}+t_kz_k$, 使 $f(x_k)\geqslant f(\bar{x})$. 由 4.3.5, $\exists\ y_k\in[\bar{x},x_k]: f(x_k)-f(\bar{x})\in\langle\partial f(y_k),t_kz_k\rangle$, 因此 $f^\circ(y_k,z_k)=\max\langle\partial f(y_k),z_k\rangle\geqslant 0$; 利用 $f^\circ(\cdot,\cdot)$usc 得出 $f^\circ(\bar{x},z)\geqslant 0$, 可见 $z\overline{\in}K_0$. 这表明 $K_0\subset F_0$. 令 $G_i=g_i^{-1}(\mathrm{R}_-)$, 则由 4.6.3 有 $K_i\subset F(G_i,\bar{x})(1\leqslant i\leqslant m)$. 以上事实结合 5.1.2 得出 $\bigcap_0^n K_j=\varnothing$, 于是由 5.1.1 有不全为零的 $u_j\in K_j^*(0\leqslant j\leqslant n): \sum u_j=0$. 由 4.6.3 之证知 $K_0^*=\mathrm{R}_-\partial f(\bar{x})$, $K_i^*=\mathrm{R}_-\partial g_i(\bar{x})(1\leqslant i\leqslant m)$; 其次有 $K_{m+1}^*=(D-\bar{x})^*$, $K_n^*=R(B^*)$. 于是 $u_0=-\bar{\rho}_0\bar{u}_0$, $u_i=-\bar{\lambda}_i\bar{u}_i$, $\bar{\rho}_0,\bar{\lambda}_i\in\mathrm{R}_+$, $\bar{u}_0\in\partial f(\bar{x})$, $\bar{u}_i\in\partial g_i(\bar{x})(1\leqslant i\leqslant m)$, $u_n=-B^*\bar{\mu}$, $\bar{\mu}\in Z^*$; $u_{m+1}=-\sum\limits_{j\neq m+1}u_j=\bar{\rho}_0\bar{u}_0+\sum\bar{\lambda}_i\bar{u}_i+B^*\bar{\mu}\in(D-\bar{x})^*$, 这得出(2). 令 $\bar{\lambda}=(\bar{\lambda}_i)$, 显然 $(\bar{\rho},\bar{\lambda},\bar{\mu})\neq 0$.

下面设 $\bar{x}\in D^\circ$, $R(B)=Z$, $\bar{\rho}=0$, 则 $(\bar{\lambda},\bar{\mu})\neq 0$, (2)推出 $\sum\bar{\lambda}_i\bar{u}_i+B^*\bar{\mu}=0$. 必定 $\bar{\lambda}\neq 0$(否则 $0\neq\bar{\mu}\in N(B^*)$, 与 $R(B)=Z$ 矛盾). 若条件(i)满足, 则有 $z\in N(B)$; $\langle\bar{u}_i,z\rangle\leqslant g^\circ_i(\bar{x},z)<0$, 于是

$$0=\langle\sum\bar{\lambda}_i\bar{u}_i+B^*\bar{\mu},z\rangle=\sum\bar{\lambda}_i\langle\bar{u}_i,z\rangle<0,$$

得出矛盾. 若条件(ii)满足, 则 $z=\hat{x}-\bar{x}$ 满足 $Bz=0$; 由 $g_i(\hat{x})<g_i(\bar{x})=0$ 与 $g_i$ 广义伪凸得 $g^\circ_i(\bar{x},z)<0$, 这同样得出矛盾. 因此 $\bar{\rho}\neq 0$. □

下面的结果所依据的思路颇不相同.

**5.4.2 定理** 设 $Z=\mathrm{R}^p$, $D$ 闭. 若 $\bar{x}$ 是(1)的局部最优解, $k>0$ 适当大, 则存在 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in\mathrm{R}_+\times\mathrm{R}_+^I\times\mathrm{R}^p$, 使得

$$0\in\partial_xL(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})+k\partial d_D(\bar{x}). \tag{3}$$

**注** 令 $\bar{\lambda}=(\bar{\lambda}_i),\bar{\mu}=(\bar{\mu}_j)$，则(3)推出

$$0\in\bar{\rho}\partial f(\bar{x})+\sum\bar{\lambda}_i\partial g_i(\bar{x})+\sum\bar{\mu}_j\partial h_j(\bar{x})+k\partial d_D(\bar{x}).\quad(4)$$

其次，由 $(D-\bar{x})^*=-\partial\delta_D(\bar{x})\subset\mathrm{R}_-\partial d_D(\bar{x})$(4.6.7)与(2)推出

$$0\in\bar{\rho}\partial f(\bar{x})+\sum\bar{\lambda}_i\partial g_i(\bar{x})+\partial(\bar{\mu}h)(\bar{x})+\mathrm{R}_+\partial d_D(\bar{x}).\quad(5)$$

这显示出 5.4.1 与 5.4.2 之间的类似.

**证** 不妨设 $f(\bar{x})=0,g(\bar{x})=0$. 取定充分小的 $\varepsilon>0$，令 $f_\varepsilon=f+\varepsilon,\omega=(f_\varepsilon,g,h),T=(\mathrm{R}_+\times\mathrm{R}_+^m\times\mathrm{R}^p)\cap S^{m+p}$. $\forall\ t=(\rho,\lambda,\mu)\in T$，定义

$$F_t(x)=\langle t,\omega(x)\rangle=L(x,t)+\rho\varepsilon;\quad(6)$$

$F(x)=\max\limits_{t\in T}F_t(x)$. 因问题是局部的，不妨设 $\mathrm{Lip}\omega<k$，于是 $\mathrm{Lip}F_t<k$，从而 $\mathrm{Lip}F<k$.

给定 $x\in D$，有 $t=(\rho,\lambda_1,\cdots,\lambda_m,\mu_1,\cdots,\mu_p)\in T$，使得

$$F(x)=F_t(x)+\rho f_\varepsilon(x)+\sum\lambda_i g_i(x)+\sum\mu_j h_j(x).\quad(7)$$

因 $\bar{x}$ 是最优解，一个细致的讨论得出：(7)右端 $1+m+p$ 项皆非负. 因此 $\rho\neq0\Rightarrow f_\varepsilon(x)\geqslant0,\lambda_i\neq0\Rightarrow g_i(x)\geqslant0$. 令 $\varphi=(f_\varepsilon^+,g_1^+,\cdots,g_m^+,h)$（约定 $f^+=\max\{f,0\}$），则前面的分析表明

$$F(x)=\langle t,\varphi(x)\rangle=|\varphi(x)|>0,x\in D,\quad(8)$$

$t=\varphi(x)/|\varphi(x)|$ 由 $x$ 唯一确定.

直接看出 $F(\bar{x})=\varepsilon\leqslant\inf\limits_D F(x)+\varepsilon$，于是由后面补证的 5.4.3（取 $\delta=\sqrt{\varepsilon}$）有 $x_\varepsilon\in D\cap\bar{B}(\bar{x},\sqrt{\varepsilon})$，使得

$$F(x_\varepsilon)=\min_{x\in D}\left[F(x)+\sqrt{\varepsilon}|x-x_\varepsilon|\right].\quad(9)$$

由 $\varepsilon$ 充分小，可设 $\mathrm{Lip}(F+\sqrt{\varepsilon}|\cdot-x_\varepsilon|)\leqslant k$，于是由(9)、4.3.7 及 4.2.2 之 3°有

$$0\in\partial(F+kd_D)(x_\varepsilon)+\sqrt{\varepsilon}B^*,\quad(10)$$

其中 $B^*=\bar{B}_1(0)\subset X^*$. 令 $R_t=F_t+kd_D,S_t=L(\cdot,t)+kd_D,R=F+kd_D$. $\forall\ t,\tau\in T$，由(6)看出

$$\mathrm{Lip}(F_t-F_\tau)=\mathrm{Lip}(R_t-R_\tau)\leqslant k|t-\tau|,$$

因此 $\partial R_t(x)\subset\partial R_\tau(x)+k|t-\tau|B^*$（4.3.2(i)），这一事实结合 4.3.2(ii)得出 $\partial_T R_t(x)=\partial R_t(x)$，记号依 4.5.1. 对 $S_t$ 有类似结论. 因 $R(x)\triangleq F(x)+kd_D(x)=\max\limits_{t\in T}R_t(x)$，而 $\partial R_t(x)=\partial S_t(x)$，故由 4.5.1 与(10)有

$$0\in\partial S_{t_\varepsilon}(x_\varepsilon)+\sqrt{\varepsilon}B^*,\tag{11}$$

$t_\varepsilon$ 由 $x_\varepsilon$ 唯一决定.

取 $\varepsilon_n\downarrow 0$，则 $x_n\triangleq x_{\varepsilon_n}\to\bar{x}$，且不妨设 $t_n\triangleq t_{\varepsilon_n}\to\bar{t}=(\bar{\rho},\bar{\lambda},\bar{\mu})\in T$. 于是由(11)得

$$0\in\partial S_{\bar{t}}(x_n)+(k|t_n-\bar{t}|+\sqrt{\varepsilon_n})B^*,$$

这结合 4.3.2(ii)得 $0\in\partial S_{\bar{t}}(\bar{x})$，由此得出(3). □

**注** 若 $\bar{x}\in D^\circ$（特别，若 $D=X$），则(3)成为

$$0\in\partial_x L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}),\tag{12}$$

这可看作 Euler-Lagrange 方程的非光滑形式.

下面补证 5.4.2 之证中用到的

**5.4.3 定理**(Ekeland，1974[46]) 设 $D\subset X$ 闭，$f:D\to\mathrm{R}$ 为 lsc，$\alpha\triangleq\inf\limits_D f(x)>-\infty$，$\varepsilon,\delta>0$，$f(\bar{x})\leqslant\alpha+\varepsilon$. 则存在 $x\in D\cap\bar{B}_\delta(\bar{x})$，使 $f(x)\leqslant f(\bar{x})$，$\forall\ y\in D\backslash\{x\}$：$f(x)<f(y)+(\varepsilon/\delta)|x-y|$. 特别，有 $x\in D$，使 $f(x)=\min\limits_{y\in D}[f(y)+\varepsilon|x-y|]\leqslant\alpha+\varepsilon$（称这样的 $x$ 为 $f$ 的"$\varepsilon$-近似极小点"）.

**证** 设定理不真，归纳地取定 $x_n$ 如下：令 $x_0=\bar{x}$；设 $x_0,\cdots,x_n$ $(n\geqslant 0)$已取定，且满足

$$|x_i-x_{i-1}|\leqslant\frac{\delta}{\varepsilon}[f(x_{i-1})-f(x_i)],i=1,2,\cdots,n.\tag{13}$$

令 $$S_n=\left\{y\in D\backslash\{x_n\}:f(x_n)\geqslant f(y)+\frac{\varepsilon}{\delta}|y-x_n|\right\},\tag{14}$$

则 $S_n\neq\varnothing$（否则 $x=x_n$ 已满足定理要求）. 令 $\beta_n=\inf f(S_n)$，则 $f(x_n)>\beta_n\geqslant\alpha$. 取 $x_{n+1}\in S_n$，使 $f(x_{n+1})<[f(x_n)+\beta_n]/2$，则(13)亦对 $i=n+1$ 成立. 如此得序列$\{x_n\}$. 显然 $f(x_n)$单调降，设 $f(x_n)$

$\to\beta$，则 $\beta\geqslant\alpha$. 若 $n>m$，则由(13)有

$$|x_n - x_m| \leqslant (\delta/\varepsilon)[f(x_m) - f(x_n)], \tag{15}$$

可见$\{x_n\}$为 Cauchy 列，设 $x_n\to x\in D$. 由(15)推出

$$\varepsilon|x_n - \bar{x}| \leqslant \delta[f(\bar{x}) - f(x_n)] \leqslant \delta(\alpha+\varepsilon-\beta) \leqslant \varepsilon\delta,$$

因此 $|x-\bar{x}|\leqslant\delta$. 由 $f$ lsc 得 $f(x)\leqslant f(\bar{x})$，且

$$f(x)\leqslant\varliminf_n f(x_n)\leqslant f(x_m)-(\varepsilon/\delta)|x-x_m|\,(m\geqslant1). \tag{16}$$

取 $y\in D\backslash\{x\}$，使 $f(x)\geqslant f(y)+(\varepsilon/\delta)|x-y|$. $\forall\, n\geqslant1$，由(16)有$f(x_n) \geqslant f(x) + (\varepsilon/\delta)|x - x_n| \geqslant f(y) + (\varepsilon/\delta)|y - x_n|$，可见 $y\in\bigcap S_n$. 因此 $f(y)\geqslant\beta_n>2f(x_{n+1})-f(x_n)\,(n\geqslant1)$. 令 $n\to\infty$ 得 $f(y)\geqslant\beta$. 但 $\beta\geqslant f(x)>f(y)$，得出矛盾. □

**注** Ekeland 定理是现代极值理论中最有意义的成果之一，有很多重要应用(如见[46]).

最后，给出一个简单的非光滑充分条件.

**5.4.4 定理** 设存在$(\bar{\lambda},\bar{\mu})\in Q\times Z^*$，使得

$$0 \in \partial_x L(\bar{x},\bar{\lambda},\bar{\mu}); \tag{17}$$

$f$ 在 $\bar{x}$ 相对于 $M$ 广义伪凸，$\varphi\triangleq\bar{\lambda}g+\bar{\mu}h$ 凸，则 $\bar{x}$ 是(1)的最优解(本定理中不要求 $Y=\mathrm{R}^m$).

**证** 若 $\bar{x}$ 非最优解，则$\exists\ x\in M$：$f(x)<f(\bar{x})$，于是 $f^\circ(\bar{x},z)<0$，$z=x-\bar{x}$. 由(17)有 $u\in\partial f(\bar{x})$，$v\in\partial\varphi(\bar{x})$：$u+v=0$. 因 $u(z)\leqslant f^\circ(\bar{x},z)<0$，故

$$0 < v(z) \leqslant \varphi(x) - \varphi(\bar{x}) = \varphi(x) = \langle\bar{\lambda},g(x)\rangle \leqslant 0,$$

得出矛盾.

## §5 类凸性与择一定理

设 $K\subset X$ 是给定的. 首先推广凸集概念*.

**5.5.1 定义** 若$\forall\ \alpha\in(0,1)$：$\alpha K+\alpha' K\subset K$，则说 $K$ 凸；若存

---

* 这一想法是沈铁告知作者的.

在有界集 $U\subset K$，使得 $\forall\ \alpha,\varepsilon\in(0,1)$；$\alpha K+\alpha' K\subset K+\varepsilon U$，则说 $K$ 次凸；若存在有界集 $U\subset X$，使得 $\forall\ \alpha,\varepsilon\in(0,1)$，$\exists\ r\in(1-\varepsilon,1+\varepsilon)$：$\alpha K+\alpha' K\subset rK+\varepsilon U$，则说 $K$ 广义次凸。若将"$\forall\ \alpha\in(0,1)$"改为"$\exists\ \alpha\in(0,1)$"，则上述概念分别改称为"$\alpha$-近凸"、"$\alpha$-近次凸"与"$\alpha$-广义近次凸"，不必指明 $\alpha$ 时省去"$\alpha$-"。

**注** 不失一般性，可设 5.5.1 中的有界集 $U$ 为对称凸集，因而 $0\in U$，$|t|\leqslant 1\Rightarrow tU\subset U$。

**5.5.2 引理** 若 $K$ 广义近次凸，则 $\overline{K}$ 凸。

**证** 取有界集 $U\subset X$，$\alpha\in(0,1)$，使得 $\forall\ \varepsilon\in(0,1)$，$\exists\ r_\varepsilon\in(1-\varepsilon,1+\varepsilon)$；$\alpha K+\alpha' K\subset r_\varepsilon K+\varepsilon U$，则

$$\alpha\overline{K}+\alpha'\overline{K}\subset\overline{\alpha K+\alpha' K}\subset\bigcap_{0<\varepsilon<1}\overline{r_\varepsilon K+\varepsilon U}\subset\overline{K}.$$

令 $T=\{t\in J:t\overline{K}+t'\overline{K}\subset\overline{K}\}$，则 $T$ 闭，$0,1,\alpha\in T$。若 $T\neq J$，则 $\exists\ \sigma,\tau\in T$：$\sigma<\tau$，$(\sigma,\tau)\subset J\backslash T$。$\beta\triangleq\alpha\sigma+\alpha'\tau$ 满足

$$\beta\overline{K}+\beta'\overline{K}=\alpha(\sigma\overline{K}+\sigma'\overline{K})+\alpha'(\tau\overline{K}+\tau'\overline{K})\subset\overline{K},$$

故 $\beta\in T\cap(\sigma,\tau)$，得出矛盾。因此 $T=J$，即 $\overline{K}$ 凸。 □

**5.5.3 引理** 若 $K$ 近次凸，则存在有界集 $U\subset X$ 与 $J$ 的稠子集 $T$，使得 $\forall\ t\in T,\varepsilon>0$：$tK+t'K\subset K+\varepsilon U$。

**证** 取有界对称凸集 $U\subset X$，$\alpha\in(0,1)$，使得 $\forall\ \varepsilon>0$：$\alpha K+\alpha' K\subset K+\varepsilon U$。令 $T=\{t\in J:tK+t'K\subset K+\varepsilon U(\forall\ \varepsilon>0)\}$，则 $0,1,\alpha\in T$。$\forall\ t_1,t_2\in T,\varepsilon>0$，$\sigma\triangleq\alpha t_1+\alpha' t_2$ 满足

$$\begin{aligned}\sigma K+\sigma' K&\subset\alpha(t_1K+t'_1K)+\alpha'(t_2K+t'_2K)\\&\subset\alpha\left(K+\frac{\varepsilon}{2}U\right)+\alpha'\left(K+\frac{\varepsilon}{2}U\right)\\&\subset\alpha K+\alpha' K+\frac{\varepsilon}{2}U\subset K+\frac{\varepsilon}{2}U+\frac{\varepsilon}{2}U\subset K+\varepsilon U,\end{aligned}$$

故 $\sigma\in T$。可见 $T$ 为 $\alpha$-近凸，从而 $\overline{T}$ 凸，故 $\overline{T}=J$。 □

**5.5.4 定理** 若 $K$ 闭，则 $K$ 凸$\Leftrightarrow K$ 广义近次凸。若 $K$ 开，则 $K$ 凸$\Leftrightarrow K$ 近凸，$K$ 次凸$\Leftrightarrow K$ 近次凸。

**证** 前半部分直接由 5.5.2 得出。下面设 $K$ 开，证 $K$ 近次凸

$\Rightarrow K$ 次凸. 设 $U$ 如 5.5.3. 若 $K$ 非次凸，则$\exists$ $\tau\in(0,1)$，$\varepsilon>0$，$x$，$y\in K$：$z\triangleq\tau x+\tau' y\overline{\in}K+\varepsilon U$. 令 $\varphi(t)=t^{-1}(z-t'y)$，则 $\varphi(\tau)=x$. 由 5.5.3 有 $t_n\to\tau$：$t_nK+t'_nK\subset K+\varepsilon U$. 由连续性可设 $\varphi(t_n)\in K$，于是 $z=t_n\varphi(t_n)+t'_ny\in K+\varepsilon U$，得出矛盾. 因此 $K$ 次凸. 类似地可证 $K$ 近凸$\Rightarrow K$ 凸. □

给定 $g:D\to Y$，约定 $Gx=g(x)+Y_+$，显然 $R(G)=g(D)+Y_+$.

**5.5.5 定义** 若 $R(G)$为凸[次凸、广义次凸、近凸、近次凸、广义近次凸]集，则称 $g$ 为类凸[次类凸、广义次类凸、近类凸、近次类凸、广义近次类凸]函数；必要指明 $Y_+$时称上述的类凸为 $Y_+$-类凸，$Y_+$-次类凸等仿此.

直接看出以上概念与通常的凸有以下关系：

凸$\Rightarrow$类凸$\Rightarrow$次类凸$\Rightarrow$广义次类凸

$\Downarrow$ $\Downarrow$ $\Downarrow$

近类凸$\Rightarrow$近次类凸$\Rightarrow$广义近次类凸

下面指明在一定条件下上面的某些"$\Rightarrow$"可换成"$\Leftrightarrow$".

**5.5.6 定理** (i)若 $g(D)+Y_+$闭，则 $g$ 类凸$\Leftrightarrow g$ 广义近次类凸. (ii)若 $Y^\circ_+\neq\varnothing$，则 $g$ 次类凸$\Leftrightarrow g$ 近次类凸$\Leftrightarrow g(D)+Y^\circ_+$凸. (iii)若 $Y^\circ_+\neq\varnothing$，$g(D)$有界，则 $g$ 次类凸$\Leftrightarrow g$ 广义近次类凸.

**证** (i)直接由 5.5.4(i)推出.

(ii)若 $g$ 近次类凸，则有有界集 $V\subset Y$，$\alpha\in(0,1)$，使得$\forall$ $\varepsilon>0$：$\alpha R(G)+\alpha' R(G)\subset R(G)+\varepsilon V$. 设 $y_i\in g(D)$，$z_i\in Y^\circ_+$ $(i=1,2)$，$\varepsilon>0$ 充分小，则

$$\alpha(y_1+z_1)+\alpha'(y_2+z_2)\in R(G)+\varepsilon V+\alpha z_1+\alpha' z_2$$
$$\subset R(G)+Y^\circ_+\subset g(D)+Y^\circ_+,$$

可见 $g(D)+Y^\circ_+$近凸，从而凸(5.5.4). 其次设 $g(D)+Y^\circ_+$凸，取定 $y_0\in Y^\circ_+$，令 $V=\{-y_0\}$. $\forall$ $t\in J$，$\varepsilon>0$，有

$$tR(G)+t'R(G)=t(R(G)+\varepsilon y_0)+t'(R(G)+\varepsilon y_0)-\varepsilon y_0$$
$$\subset t(g(D)+Y^\circ_+)+t'(g(D)+Y^\circ_+)+\varepsilon V$$

$$\subset g(D) + Y^{\circ}_{+} + \varepsilon V \subset R(G) + \varepsilon V,$$

可见 $R(G)$次凸，从而 $g$ 次类凸.

(iii)若 $g$ 广义近次类凸，则有有界对称凸集 $V\subset Y$，$\alpha\in(0,1)$，使得$\forall\ \varepsilon\in(0,1)$，$\exists\ r\in(1-\varepsilon,1+\varepsilon)$：$\alpha R(G)+\alpha' R(G)\subset rR(G)+\varepsilon V$，可设 $g(D)\subset V$，于是

$$\alpha R(G) + \alpha' R(G) \subset R(G) + (r-1)g(D) + \varepsilon V$$
$$\subset R(G) + \varepsilon V + \varepsilon V \subset R(G) + 2\varepsilon V,$$

可见 $g$ 近次类凸，从而由已证之(ii)必次类凸. □

现在转入择一定理. 如下形式的断言称为"择一定理"："命题 $P$ 与 $Q$ 恰有一个成立"(或简言之：$P$ 与 $Q$ 两择一). 显然以上断言相当于"$P\Leftrightarrow$非 $Q$"，因此任何等价命题产生择一定理. 不过，本节所述的择一定理都涉及一定不等式系统的可解性，因而亦称为"可解性定理"，它们是最优化理论的基本工具之一.

以下设 $W$ 与 $Y$ 分别由闭凸锥 $W_+$ 与 $Y_+$ 导入序$\leqslant$，假定 $W^{\circ}_{+}\neq\varnothing$. 给定 $A_0\in L(X,W)$，$A\in L(X,Y)$，$B\in L(X,Z)$，$f:D\to W$，$g:D\to Y$，$h:D\to Z$，$L(\cdot,\rho,\lambda,\mu)=\rho f+\lambda g+\mu h$. 若 $x_0\in D$，$\forall\ \varepsilon>0$，$\exists\ \delta>0$：$g(D\cap B_\delta(x_0))\subset B_\varepsilon(g(x_0))-Y_+$，则说 $g$ 在 $x_0$(关于 $Y_+$)usc.

**5.5.7 广义 Motzkin 定理** 设以下条件满足：

($H_1$) $\forall\ t\in(0,1)$，$x_1,x_2\in D$，$\exists\ \beta>0$，$\forall\ \varepsilon\in(0,1)$，$\exists\ r,s\in(1-\varepsilon,1+\varepsilon)$，$(x,u,v)\in D\times W\times Y$：$rf(x)\leqslant tf(x_1)+t'f(x_2)+\varepsilon u$，$sg(x)\leqslant tg(x_1)+t'g(x_2)+\varepsilon v$，$h(x)=th(x_1)+t'h(x_2)$，且$|f(x)|$，$|g(x)|$，$|u|$，$|v|\leqslant\beta$；

($H_2$) (Slater 条件)$\exists\ \hat{x}\in D$：$g(\hat{x})\ll 0$，$h(\hat{x})=0$；

($H_3$) $f$ 与 $g$ 在 $\hat{x}$(依($H_2$))usc；

($H_4$) $\forall\ \delta>0$，$\exists\ \eta>0$：$B_\eta(0)\subset h(D\cap B_\delta(\hat{x}))$，

则以下命题两择一：(P)$\exists\ x\in D$：$f(x)\ll 0$，$g(x)\leqslant 0$，$h(x)=0$；(Q)$\exists\ (\rho,\lambda,\mu)\in(W^*_+\setminus\{0\})\times Y^*_+\times Z^*$：$L(D,\rho,\lambda,\mu)\geqslant 0$. 若 $W$，$Y$，$Z$ 有限维，条件($H_1$)中限定 $s=1$，$v=0$，则($H_2$)～($H_4$)可代以条

件

($H_5$) $\{(\lambda,\mu)\in Y_+^*\times Z^*:(\lambda g+\mu h)(D)\geqslant 0\}=\{0\}$.

**证** (P)与(Q)显然互斥，下面证(P)无解⇒(Q)可解. 令 $K=W^\circ_+\times Y^\circ_+\times\{0\}$，$\varphi=(f,g,h)$，$C=\varphi(D)+K$. $\forall\, t\in(0,1)$，$x_i\in D$，$w_i\in W^\circ_+$，$y_i\in Y^\circ_+(i=1,2)$，有 $w=tw_1+t'w_2\in W^\circ_+$，$y=ty_1+t'y_2\in Y^\circ_+$. 设 $\beta,r,s,x,u,v$ 如条件($H_1$)，$\varepsilon$ 充分小，则

$$\begin{aligned}&t[\varphi(x_1)+(w_1,y_1,0)]+t'[\varphi(x_2)+(w_2,y_2,0)]\\&=(tf(x_1)+t'f(x_2)+w,tg(x_1)+t'g(x_2)+y,h(x))\\&\in\varphi(x)+(w+(r-1)f(x)-\varepsilon u,y\\&\quad+(s-1)g(x)-\varepsilon v,0)+W_+\times Y_+\times\{0\}\\&\subset\varphi(D)+K+W_+\times Y_+\times\{0\}\subset C,\end{aligned}$$

可见 $C$ 凸. 取 $\overline{w}\gg f(\hat{x})$，设 $(w,y,z)$ 邻近 $(\overline{w},0,0)$. 取 $\varepsilon>0$，使 $B_\varepsilon(f(\hat{x}))\ll w$，$B_\varepsilon(g(\hat{x}))\ll y$. 由条件($H_3$)有 $\delta>0$：$f(D\cap B_\delta(\hat{x}))\subset B_\varepsilon(f(\hat{x}))-W_+$，$g(D\cap B_\delta(\hat{x}))\subset B_\varepsilon(g(\hat{x}))-Y_+$. 由($H_4$)，可设 $\exists\, x\in B_\delta(\hat{x})\cap h^{-1}(z)$. 易见 $w-f(x)\in W^\circ_+$；同理 $y-g(x)\in Y^\circ_+$，于是 $(w,y,z)=\varphi(x)+(w-f(x),y-g(x),0)\in C$，可见 $(\overline{w},0,0)\in C^\circ$. (P)无解⇒$0\overline{\in}C$，于是由分离定理有 $0\neq\omega=(\rho,\lambda,\mu)\in(W\times Y\times Z)^*$：$0\leqslant\langle\omega,C\rangle=\langle\omega,\varphi(D)\rangle+\langle\omega,K\rangle$，这推出 $L(D,\rho,\lambda,\mu)=\langle\omega,\varphi(D)\rangle\geqslant 0$，$\omega\in K^*=W_+^*\times Y_+^*\times Z^*$. 必 $\rho\neq 0$，否则 $(\lambda g+\mu h)(\hat{x})\geqslant 0$，这推出 $\lambda=0$，$0\neq\mu\in h(D)^*$，与条件($H_4$)矛盾. 因此 $(\rho,\lambda,\mu)$ 满足(Q).

若 $W,Y,Z$ 有限维，则改令 $C=\varphi(D)+W^\circ_+\times Y_+\times\{0\}$，无需 $C^\circ\neq\varnothing$ 亦可用分离定理得所要证. □

若 $(f,g,h)$ 为 $W_+\times Y_+\times\{0\}$-类凸，则 5.5.7 之条件($H_1$)显然满足. 据此易从 5.5.7 得出：

**5.5.8 推论**(线性 Motzkin 定理) 若$\exists\,\hat{x}\in N(B)$：$A\hat{x}\ll 0$ 且 $R(B)=Z$(或 $W,Y,Z$ 有限维且 $\{(\lambda,\mu)\in Y_+^*\times Z^*:A^*\lambda+B^*\mu=0\}=\{0\}$)，则以下命题两择一：(P)$\exists\, x\in N(B)$：$A_0x\ll 0$，$Ax\ll 0$；(Q)

$\exists\ (\rho,\lambda,\mu)\in(W_+^*\setminus\{0\})\times Y_+^*\times Z^*: A_0^*\rho+A^*\lambda+B^*\mu=0.$

**5.5.9 定理** 设 $Y^\circ_+\neq\varnothing$, $Y_+^*\cap g(D)^*=\{0\}$. 若 $\varphi\triangleq(f,g)$关于 $W_+\times Y_+$为近次类凸或广义近次类凸且 $\varphi(D)$有界，则以下命题两择一：(P)$\exists\ x\in D: f(x)\ll 0, g(x)\leqslant 0$；(Q)$\exists\ (\rho,\lambda)\in(W_+^*\setminus\{0\})\times Y_+^*: (\rho f+\lambda g)(D)\geqslant 0.$

**证** 由 5.5.6, $C\triangleq\varphi(D)+W^\circ_+\times Y^\circ_+$凸，于是如同证 5.5.7 一样可用分离定理得所要证. □

在 5.5.9 中取 $Y=\{0\}$得出：

**5.5.10 推论(非线性 Gordan 定理)** 若 $f$ 近次类凸或广义近次类凸且 $f(D)$有界，则以下命题两择一：(P)$\exists\ x\in D: f(x)\ll 0$；(Q)$\exists\ \rho\in W_+^*\setminus\{0\}: \langle\rho, f(D)\rangle\geqslant 0.$

**5.5.11 推论(线性 Gordan 定理)** 以下命题两择一：(P)$\exists\ x\in X: A_0x\ll 0$；(Q)$\exists\ \rho\in W_+^*\setminus\{0\}: A_0^*\rho=0.$

若在 5.5.7 之证中 $C$ 闭，则不必验证 $C^\circ\neq\varnothing$ 即可用分离定理，且有 $\langle\omega,\varphi(D)\rangle>0$. 因此，在 5.5.7 中取 $W=\{0\}$且在其证中改令 $C=\varphi(D)+Y_+\times\{0\}$得到：

**5.5.12 推论** 若 $\varphi\triangleq(g,h)$为 $Y_+\times\{0\}$-广义近次类凸且 $\varphi(D)+Y_+\times\{0\}$闭，则以下命题两择一：(P)$\exists\ x\in D: g(x)\leqslant 0, h(x)=0$；(Q)$\exists\ (\lambda,\mu)\in Y_+^*\times Z^*: (\lambda g+\mu h)(D)>0.$

在 5.5.12 中取 $Z=\{0\}$得到：

**5.5.13 推论(Gale 定理)** 若 $g$ 为 $Y_+$-广义近次类凸，$g(D)+Y_+$闭，则以下命题两择一：(P)$\exists\ x\in D: g(x)\leqslant 0$；(Q)$\exists\ \lambda\in Y_+^*: \langle\lambda, g(D)\rangle>0$. 特别，若 $R(A)+Y_+$闭(当 $Y=\mathrm{R}^m$ 时必定如此)，$a\in Y$，则“$\exists\ x\in X: Ax\leqslant a$”与“$\exists\ \lambda\in Y_+^*: A^*\lambda=0$ 且 $\lambda(a)=-1$”两择一.

**注** 由 5.5.6, 5.5.12 与 5.5.13 中的“广义近次类凸”实际上是“类凸”，表为弱形式意在便于检验.

今考虑将择一定理用于问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in X. \tag{1}$$

设 $\bar{x} \in M = g^{-1}(Y_-) \cap h^{-1}(0)$, $f, g, h$ 在 $\bar{x}$ 可微, $\tilde{f}(x) = f(x) - f(\bar{x})$, $L(\cdot, \lambda, \mu) = f + \lambda g + \mu h$. 显然 $\bar{x}$ 是(1)的最优解$\Leftrightarrow$问题"$\exists\, x \in X$: $\tilde{f}(x) < 0, g(x) \leqslant 0, h(x) = 0$"无解，这正好应用 5.5.7.

**5.5.14 定理** 设 $\bar{x}$ 是(1)的最优解，$(f, g, h)$满足 5.5.7 中条件$(H_1)$～$(H_4)$(取 $W = \mathbf{R}$)，则 $\bar{x}$ 是 $K$-$T$ 点.

**证** 易见$(\tilde{f}, g, h)$亦满足 5.5.7 中条件$(H_1)$～$(H_4)$，于是由 5.5.7 有$(\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in (0, \infty) \times Y_+^* \times Z^*$，使得

$$\bar{\rho}[f(x) - f(\bar{x})] + \bar{\lambda} g(x) + \bar{\mu} h(x) \geqslant 0 (\forall\, x \in X). \tag{2}$$

不妨设 $\bar{\rho} = 1$(否则以 $\bar{\lambda}/\bar{\rho}, \bar{\mu}/\bar{\rho}$ 代 $\bar{\lambda}, \bar{\mu}$). 以 $x = \bar{x}$ 代入(2)得$\langle \bar{\lambda}, g(\bar{x}) \rangle = 0$. 于是(2)推出

$$f(\bar{x}) = L(\bar{x}, \bar{\lambda}, \bar{\mu}) = \min_{x \in X} L(x, \bar{\lambda}, \bar{\mu}),$$

因此 $L_x(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$，从而 $\bar{x}$ 是 $K$-$T$ 点. □

$\bar{x}$ 是(2)的最优解并相当于问题"$\exists\, x \in h^{-1}(0)$: $\tilde{f}(x) < 0$, $g(x) \leqslant 0$"无解，这导致应用 5.5.9.

**5.5.15 定理** 设 $\bar{x}$ 是(1)的最优解，$h(x) = Bx - b$, $R(B)$闭；$D = h^{-1}(0)$, $(f, g)|D$ 关于 $\mathbf{R}_+ \times Y_+$ 为近次类凸或广义近次类凸且 $f(D), g(D)$有界；$\exists\, \hat{x} \in M$: $g(\hat{x}) \ll 0$，则 $\bar{x}$ 是 $K$-$T$ 点.

**证** 对$(\tilde{f}, g)|D$ 应用 5.5.9 得$(\bar{\rho}, \bar{\lambda}) \in (0, \infty) \times Y_+^*$，使得

$$\bar{\rho}[f(x) - f(\bar{x})] + \langle \bar{\lambda}, g(x) \rangle \geqslant 0 (\forall\, x \in D). \tag{3}$$

如同证 5.5.14，有$\langle \bar{\lambda}, g(\bar{x}) \rangle = 0$，且可设 $\bar{\rho} = 1$，于是(3)推出 $f(\bar{x}) = (f + \bar{\lambda} g)(\bar{x}) = \min(f + \bar{\lambda} g)(D)$. 因 $D$ 凸，故

$$f'(\bar{x}) + \bar{\lambda} \circ g'(\bar{x}) \in (D - \bar{x})^* = N(B) = R(B^*)$$

(用 4.1.10(iv)与闭值域定理)，由此显然得所要证. □

一个平行于 5.5.15 的非光滑结果是：

**5.5.16 定理** 设 $\bar{x}$ 是(1)的最优解，$h(x) = Bx - b$, $R(B)$闭；$D = h^{-1}(0)$, $(f, g)$在 $\bar{x}$ 邻近为 Lip, $(f, g)|D$ 关于 $\mathbf{R}_+ \times Y_+$ 为近次类凸或广义近次类凸且 $f(D), g(D)$有界；$\exists\, \hat{x} \in M$: $g(\hat{x}) \ll 0$,

则$\exists$ $(\bar\lambda,\bar\mu)\in Y_+^*\times Z^*$，使得$\langle\bar\lambda,g(\bar x)\rangle=0$，且

$$0\in\partial f(\bar x)+\partial(\bar\lambda g)(\bar x)+B^*\bar\mu. \tag{4}$$

**证** 用5.5.15的证法，只是注意用4.3.8从$f(\bar x)=\min(f+\bar\lambda g)(D)$得出$0\in\partial(f+\bar\lambda g)(\bar x)+\partial\delta_D(\bar x)$，后者推出(4). □

## §6 Minimax定理与鞍点

以下四节给出非线性最优化的对偶理论. 对偶理论的一般提法是：给定"原问题"

$$\min f(x)=\alpha, x\in D \tag{P}$$

($\alpha$表最优值)，构成一个与之关联的"对偶问题"

$$\max F(y)=\beta, y\in K, \tag{P*}$$

使得关于二者的论断有某种肯定的关系；特别，我们希望有$\alpha=\beta$(强对偶性)，或至少$\alpha\geqslant\beta$(弱对偶性). 若$\bar x$与$\bar y$分别为(P)与(P*)的解且$\alpha=\beta$，则它们满足所谓"极值关系"

$$f(\bar x)=F(\bar y), \tag{1}$$

对于由$\bar x$确定$\bar y$或由$\bar y$确定$\bar x$，(1)可能是有用的.

有多种途径形成对偶问题. 一种典型方式是，取一[illegible] $L:D\times K\to\mathrm{R}$，使得

$$f(x)=\sup L(x,K), F(y)=\inf L(D,y). \tag{2}$$

令
$$\alpha=\inf_x\sup_y L(x,y), \beta=\sup_y\inf_x L(x,y), \tag{3}$$

则$\alpha\geqslant\beta$平凡地成立，而$\alpha=\beta$是不常有的. 当$\alpha=\beta$时可能有三种情况，即

$$\min_x\sup_y L(x,y)=\max_y\inf_x L(x,y); \tag{4}$$

$$\min_x\sup_y L(x,y)=\sup_y\inf_x L(x,y); \tag{5}$$

$$\inf_x\sup_y L(x,y)=\max_y\inf_x L(x,y). \tag{6}$$

(4)$\Leftrightarrow$(P)(P*)皆有解且$\alpha=\beta$；(5)$\Leftrightarrow$(P)有解且$\alpha=\beta$；(6)$\Leftrightarrow$(P*)有解且$\alpha=\beta$. 断言(4)～(6)中某个成立的任何结果称为"mini

max 定理”，本书将建立几个这样的定理，它们对于最优化理论是重要的.

下面设已给定函数 $L:D\times K\to\mathrm{R}$, $f$, $F$, $\alpha$, $\beta$ 依(2)(3). 若 $L(\bar{x},\bar{y})=\alpha=\beta$，则称 $(\bar{x},\bar{y})$ 为 $L$ 的鞍点. 若 $\forall\ (x,y)\in D\times K$: $L(\cdot\ ,y)$ 与 $-L(x,\cdot\ )$ 皆[拟]凸，则称 $L$ 为[拟]鞍函数; $L$ 半连续 $\Leftrightarrow$ $L(x,y)$ 对 $x$ 为 lsc 对 $y$ 为 usc.

**5.6.1 命题** $(\bar{x},\bar{y})$ 是 $L$ 的鞍点 $\Leftrightarrow L(\bar{x},K)\leqslant L(\bar{x},\bar{y})\leqslant L(D,\bar{y})\Leftrightarrow\bar{x}$ 与 $\bar{y}$ 分别为(P)与($\mathrm{P}^*$)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow f(\bar{x})=F(\bar{y})(=L(\bar{x},\bar{y}))$.

证明是平凡的. 注意 5.6.1 无需关于 $L$ 与 $D,K$ 的任何特殊假定.

5.6.1 表明，一旦求出了 $L$ 的某个鞍点，就同时得到了(P)与($\mathrm{P}'$)的解. 然而，鞍点的出现是极不寻常的. 下面给出一个鞍点存在定理.

**5.6.2 定理** 设 $D\subset X$, $K\subset Y$, $X$ 与 $Y$ 自反，$D$ 与 $K$ 闭凸，$L$ 是半连续鞍函数. 给定条件:

(H) $\exists\ \hat{y}\in K$: $|x|\to\infty(x\in D)\Rightarrow L(x,\hat{y})\to\infty$;

($\mathrm{H}^*$) $\exists\ \hat{x}\in D$: $|y|\to\infty(y\in K)\Rightarrow L(\hat{x},y)\to-\infty$.

(i)若(H)($\mathrm{H}^*$)满足，则 $L$ 有鞍点. (ii)若(H)满足且 $\alpha<\infty$，则(5)成立. (iii)若($\mathrm{H}^*$)满足且 $\beta>-\infty$，则(6)成立.

**证** (i)首先设 $D,K$ 有界. 因 $X$ 自反，不妨设 $X$ 严格凸(3.2.4)，进而可设 $L(\cdot\ ,y)$ 严格凸(不然以 $L(x,y)+\varepsilon|x|$ 代 $L(x,y)$ 然后令 $\varepsilon\downarrow 0$)，于是 $\forall\ y\in K$，有唯一 $x=\varphi(y)\in D$: $F(y)=L(\varphi(y),y)$(4.1.10). 直接验知 $F$ 凹且 usc，于是 $\exists\ y_0\in K$: $F(y_0)=\beta$. $\forall\ y\in K$，令 $y_t=ty+t'y_0(0<t<1)$, $x_t=\varphi(y_t)$，则 $F(y_t)=L(x_t,y_t)$,

$$F(y_0)\geqslant L(x_t,y_t)\geqslant tL(x_t,y)+t'L(x_t,y_0)$$
$$\geqslant tL(x_t,y)+t'F(y_0);\tag{7}$$
$$t'L(x_t,y_0)\leqslant F(y_0)-tL(x_t,y)\leqslant F(y_0)-tF(y).\tag{8}$$

取 $t_n \downarrow 0$，使 $x_n \triangleq x_{t_n} \to x_0 \in D$. 由 $L(\cdot, y)$wlsc(4.1.6)及(8)有

$$F(y_0) \leqslant L(x_0, y_0) \leqslant \underline{\lim_n} L(x_n, y_0) \leqslant F(y_0),$$

可见 $F(y_0) = L(x_0, y_0)$，$x_0 = \varphi(y_0)$. (7)推出 $L(x_t, y) \leqslant F(y_0)$，于是

$$L(x_0, y) \leqslant \underline{\lim_n} L(x_n, y) \leqslant F(y_0) (\forall\ y \in K),$$

因此 $\alpha \leqslant f(x_0) \leqslant F(y_0) = \beta$，可见 $(x_0, y_0)$ 为鞍点.

对一般情况，令 $D_n = \{x \in D : |x| \leqslant n\}$，$K_n$ 仿此. 由已证结论，当 $n$ 充分大时 $L|D_n \times K_n$ 有鞍点 $(x_n, y_n)$，即

$$L(x_n, K_n) \leqslant L(x_n, y_n) \leqslant L(D_n, y_n). \tag{9}$$

(9)结合条件(H)(H$^*$)推出 $x_n, y_n$ 与 $L(x_n, y_n)$ 有界，因此不妨设 $x_n \to \bar{x}$，$y_n \to \bar{y}$，$L(x_n, y_n) \to r$. 由(9)推出 $L(\bar{x}, K) \leqslant r \leqslant L(D, \bar{y})$，由此易见 $(\bar{x}, \bar{y})$ 为 $L$ 之鞍点.

(ii)令 $L_n(x, y) = L(x, y) - n^{-1}|y|^2 (n \geqslant 1)$，则 $L_n$ 如同 $L$ 一样满足条件(H)；利用 4.1.7 可指明 $L_n$ 亦满足条件(H$^*$)，于是由已证之(i)知 $L_n$ 有鞍点 $(x_n, y_n)$，因此

$$L(x_n, y) - n^{-1}|y|^2 \leqslant L(x_n, y_n) - n^{-1}|y_n|^2$$

$$\leqslant L(x_n, y_n) \leqslant L(D, y_n) (\forall\ y \in K). \tag{10}$$

(10)推出 $L(x_n, y) - n^{-1}|y|^2 \leqslant \beta \leqslant \alpha (\forall\ y \in K)$；用条件(H)得出 $x_n$ 有界. 不妨设 $x_n \to \bar{x}$，则 $L(\bar{x}, K) \leqslant \beta$，这推出(5).

(iii)的证明是类似的. □

下面继续考虑使(6)成立的条件(注意(5)与(6)是类似的)，一方面以某种紧性条件代替 5.6.2 中的条件(H$^*$)，另一方面放宽凸性要求. 约定以下术语：若 $\forall\ \alpha \in (0,1)$，$x_1, x_2 \in D$，$\exists\ x \in D$：$L(x, \cdot) \leqslant \alpha L(x_1, \cdot) + \alpha' L(x_2, \cdot)$，则说 $L$ 在 $D$ 上类凸；若 $\forall\ \alpha, \varepsilon \in (0,1)$，$x_1, x_2 \in D$，$\exists\ x \in D$：$L(x, \cdot) \leqslant \alpha L(x_1, \cdot) + \alpha' L(x_2, \cdot) + \varepsilon$，则说 $L$ 在 $D$ 上次类凸；若 $\forall\ \alpha, \varepsilon \in (0,1)$，$\exists\ r \in (1-\varepsilon, 1+\varepsilon)$，$\forall\ x_1, x_2 \in D$，$\exists\ x \in D$：$rL(x, \cdot) \leqslant \alpha L(x_1, \cdot) + \alpha' L(x_2, \cdot) + \varepsilon$，则说 $L$ 在 $D$ 上广义次类凸. 参照 5.5.1，$L$ 在 $D$ 上"近类凸"、"近次

类凸"与"广义近次类凸"的意义自明. 若$-L$在$K$上类凸，则说$L$在$K$上类凹；次类凹等仿此. 约定$\omega$恒记$D$的有限子集，令$K(\omega,\sigma)=\{y:L(\omega,y)\geqslant\sigma\}(\sigma\in\mathrm{R})$. 当$L(x,\cdot)(\forall x\in D)$usc 时$K(\omega,\sigma)$为闭集.

**5.6.3 定理** 设$K$是一 Hausdorff 空间，存在$\omega_0\subset D,\sigma_0\in(-\infty,\alpha):K_0\triangleq K(\omega_0,\sigma_0)$紧；$\forall x\in D:L(x,\cdot)$usc，则以下条件之一推出(6)成立：

($C_1$) $L$在$D$上次类凸，在$K$上近次类凹；

($C_2$) $L$在$D$上近次类凸，在$K$上广义近次类凹，$\forall x\in D$：$L(x,K)$有界.

**证** $\Gamma\triangleq\{K(\omega,\sigma):\omega_0\subset\omega\subset D,\sigma_0\leqslant\sigma<\alpha\}$是$K_0$中的闭集族. 若$\varnothing\bar{\in}\Gamma$，则由$K(\bigcup\omega_i,\max\sigma_i)\subset\bigcap K(\omega_i,\sigma_i)$知$\Gamma$有限相交，从而由$K_0$紧得出$\Gamma$中诸集有公共点$\bar{y}$. $\forall x\in D$，令$\omega_x=\omega_0\cup\{x\}$，$\forall\sigma\in(\sigma_0,\alpha)$，由$\bar{y}\in K(\omega_x,\sigma)$得$L(x,\bar{y})\geqslant\sigma$. 这推出$L(D,\bar{y})\geqslant\alpha$，由此显然得出(6). 余下只要证$\varnothing\bar{\in}\Gamma$.

假定$\varnothing\in\Gamma$，于是有$\omega=\{x_1,\cdots,x_n\}\subset D,\sigma<\alpha:K(\omega,\sigma)=\varnothing$. 令$g_i(y)=\sigma-L(x_i,y)(1\leqslant i\leqslant n)$，$g=(g_i)$，今对$g:K\to\mathrm{R}^n$用 Gordan 定理(5.5.10). 条件($C_1$)推出$g$近次类凸；条件($C_2$)推出$g$广义近次类凸且$g(K)$有界；$K(\omega,\sigma)=\varnothing$推出问题"$\exists y\in K$：$g(y)\ll 0$"无解，于是有$0\neq\lambda=(\lambda_i)\in\mathrm{R}^n_+:\langle\lambda',g(K)\rangle\geqslant 0$. 可设$\sum\lambda_i=1$，于是$\sum\lambda_iL(x_i,y)\leqslant\sigma(\forall y\in K)$. 若$L$在$D$上次类凸，$\varepsilon>0$，则对$n$用归纳法可得出$x_\varepsilon\in D$：

$$L(x_\varepsilon,y)\leqslant\sum\lambda_iL(x_i,y)+\varepsilon\leqslant\sigma+\varepsilon(\forall y\in K).$$

这推出$\alpha\leqslant f(x_\varepsilon)\leqslant\sigma+\varepsilon$；令$\varepsilon\to 0$得$\alpha\leqslant\sigma$，得出矛盾.

余下只要证($C_2$)$\Rightarrow L$在$D$上次类凸. 设($C_2$)满足. 因$L$在$D$上近次类凸，如同证 5.5.3 一样可指明存在$J$的稠子集$T$，使得$\forall t\in T,\forall\varepsilon>0,x_1,x_2\in D,\exists x\in D$：

$$L(x,\cdot)\leqslant tL(x_1,\cdot)+t'L(x_2,\cdot)+\varepsilon. \tag{11}$$

若 $L$ 在 $D$ 上非次类凸，则 $\exists\ t_0\in(0,1),\varepsilon_0>0,x_1,x_2\in D$，使得 $\forall\ x\in D$，$\exists\ y_x\in K$：

$$L(x,y_x)>t_0L(x_1,y_x)+t'_0L(x_2,y_x)+2\varepsilon_0. \tag{12}$$

取 $t_n\to t_0$；$x_n\in D$，使(11)对 $x=x_n,t=t_n,\varepsilon=\varepsilon_0$ 成立. 令 $y_n=y_{x_n}$，则结合(11)(12)得

$$\varepsilon_0<(t_n-t_0)[L(x_1,y_n)-L(x_2,y_n)]\ (n\geqslant 1). \tag{13}$$

因 $L(x_i,y_n)(i=1,2)$ 有界，故当 $n\to\infty$ 时(13)之右端 $\to 0$，得出矛盾. □

所宜注意者，5.6.3 对 $D$ 没有任何要求. 若 $K$ 本身紧，则 5.6.3 中的紧性条件自动满足.

在 $K$ 紧的条件下建立的类似于 5.6.3 的结果颇多，今择其主要者不加证明地陈述于下.

**5.6.4 定理** 以下每个条件推出(6)成立：

($C_3$) $K$ 为紧 Hausdorff 空间，$L(x,\cdot)(\forall\ x\in D)$ usc，$L$ 在 $D$ 上近次类凸在 $K$ 上近次类凹(简称"近次类鞍")；

($C_4$) $K$ 是某拓扑向量空间(TVS)中的凸紧集，$L$ 在 $D$ 上近次类凸，$\forall\ x\in D$：$L(x,\cdot)$ 拟凹且 usc；

($C_5$) $D$ 是某 TVS 中的凸集，$K$ 是某 TVS 中的凸紧集，$L$ 拟鞍、半连续.

minimax 定理是择一定理的重要来源之一. 下面说明由(6)得出择一定理的一般模式. 设 $L:D\times Y_+\to\mathrm{R}$，$L(x,y)$ 对 $y$ 是正齐次的，$K$ 是 $Y_+$ 的基，这意味着 $0\overline{\in}K\subset Y_+=\mathrm{R}_+K$. 若 $L|D\times K$ 满足(6)，则显然以下命题两择一：

$$\exists\ x\in D:L(x,Y_+\setminus\{0\})<0; \tag{14}$$

$$\exists\ y\in Y_+\setminus\{0\}:L(D,y)\geqslant 0. \tag{15}$$

因此，任何断言(6)成立的结果导致择一定理. 特别，从 5.6.3 与 5.6.4 得出：

**5.6.5 定理** 设 $Y$ 为 TVS，$K$ 是 $Y_+$ 的基，$L:D\times Y_+\to\mathrm{R}$，$\forall\ x\in D$：$L(x,\cdot)$ 正齐次且 usc，则以下每个条件推出(14)与(15)两

择一：(i)5.6.3之条件满足；(ii)$K$紧，$L|D\times K$近次类鞍，(iii)$K$紧凸，$L$在$D$上近次类凸，$\forall x\in D: L(x,\cdot)|K$拟凹；(iv)$K$紧凸，$D\subset X$凸，$X$为TVS，$L|D\times K$拟鞍、半连续.

将5.6.5用于$L(x,\lambda)=\langle\lambda,g(x)\rangle$(以$Y_+^*$代$Y_+$)得出：

**5.6.6推论** 设$g: D\subset X\to Y$，$K$是$Y_+^*$的凸基. 假定以下条件之一满足：(i)$\exists\ \omega\subset D$，$-\infty<\sigma<\inf_x \sup\langle K,g(x)\rangle$：$\{\lambda\in K: \langle\lambda,g(\omega)\rangle\geqslant\sigma\}$弱$^*$紧，$g$类凸(或$g$近次类凸而$K$有界)；(ii)$K$弱$^*$紧，$g$近次类凸；(iii)$K$弱$^*$紧，$g$为$*$拟凸且$*$lsc(依4.1.11)，则以下命题两择一：(P)$\exists\ x\in D: \langle Y_+^*\backslash\{0\},g(x)\rangle<0$；(Q)$\exists\ \lambda\in Y_+^*\backslash\{0\}: \langle\lambda,g(D)\rangle\geqslant 0$.

**证** 令$L(x,\lambda)=\langle\lambda,g(x)\rangle$，$(x,\lambda)\in D\times Y_+^*$，在$Y^*$中采用弱$^*$拓扑. $L(x,\cdot)$显然是齐次、连续的. 当$K$弱$^*$紧时$K$必有界，用此易从$g$近次类凸推出$L|D\times K$在$D$上近次类凸. 由$K$凸推出$L(x,\lambda)$对$\lambda\in K$凸. 基于此，易见5.6.6之条件(i)(ii)(iii)分别推出5.6.5之条件(i)(ii)(iv)，于是所要结论由5.6.5推出. □

**注** 5.6.6可看作是一个"Gordan型定理". 与5.5.10比较，5.6.6的主要好处是无需假定$Y^\circ_+\neq\varnothing$，且所用条件有更大的选择余地. 若$Y^\circ_+\neq\varnothing$，则5.6.6的结论与5.5.10完全一致. 取定$y\in Y^\circ_+$，可验知$K\triangleq\{\lambda\in Y_+^*: \lambda(y)=1\}$是$Y_+^*$的弱$^*$紧凸基. 因此当5.5.10之条件满足时5.6.6之条件(ii)满足(参考5.5.6)，这表明5.6.6蕴涵5.5.10.

举一个应用5.6.6的例子. 考虑问题

$$\min f(x), g(x)\leqslant 0, Bx=b, x\in X, \tag{16}$$

其中$f,g$在$\bar{x}$可微，$\bar{x}\in D\cap g^{-1}(Y_-)$，$D=B^{-1}(b)$. 令$\mathcal{J}(x)=f(x)-f(\bar{x})$(参照5.5.15之证). 若$\bar{x}$是(16)的解，则问题

$$\exists\ x\in D: \langle(\mathrm{R}_+\times Y_+^*)\backslash\{0\},(\mathcal{J}(x),g(x))\rangle<0 \tag{17}$$

必无解. 于是自然考虑对$(\mathcal{J},g)|D$应用5.6.6. 假定$(\mathcal{J},g)|D$已满足5.6.6之条件，则由5.6.6有

$$\exists\ (\rho,\lambda)\in(\mathrm{R}_+\times Y_+^*)\backslash\{0\}:(\rho\bar{f}+\lambda g)(D)\geqslant 0. \tag{18}$$

用一个"Slater 型条件"

$$\exists\ \hat{x}\in D:\langle Y_+^*\backslash\{0\},g(\hat{x})\rangle<0 \tag{19}$$

可从(18)得出 $\rho\neq 0$，因而可设 $\rho=1$. 于是由(18)有 $f(\bar{x})\leqslant(f+\lambda g)(D)$，因而可如证 5.5.15 一样得出以下结果：

**5.6.7 定理** 设 $\bar{x}$ 是(16)的最优解，$D=B^{-1}(b)$，$R(B)$闭；$Y_+^*$ 有弱*紧凸基 $K$，条件(19)满足. 若 $(f,g)|D$ 关于锥 $\mathrm{R}_+\times Y_+$ 为近次类凸，或*拟凸且*lsc，则 $\bar{x}$ 是(16)的 $K$-$T$ 点.

**证** 只需指明以下两点：(A)不难验证 $\Gamma\triangleq\{(t,t'\lambda):(t,\lambda)\in J\times K\}$ 是 $R_+\times Y_+^*$ 的弱*紧凸基；(B)$(\bar{f},g)|D$ 如同 $(f,g)|D$ 一样，关于锥 $\mathrm{R}_+\times Y_+$ 近次类凸，或*拟凸且*lsc. 以上事实表明 5.6.6 之条件(ii)或(iii)满足. □

**注** 若 $Y^\circ_+\neq\varnothing$，则条件$(19)\Leftrightarrow\exists\ \hat{x}\in D:g(\hat{x})\ll 0$. 由此可见 5.6.7 蕴涵了 5.5.15，这与 5.6.6 蕴涵 5.5.10 是恰相呼应的.

相应的非光滑结果是(参照 5.5.16)：

**5.6.8 定理** 设 $\bar{x}$ 是(16)的最优解. 若 5.6.7 之条件满足，且 $f,g$ 在 $\bar{x}$ 邻近为 Lip(但不假定 $f,g$ 在 $\bar{x}$ 可微)，则$\exists\ (\bar{\lambda},\bar{\mu})\in Y_+^*\times Z^*$，使得 $\langle\bar{\lambda},g(\bar{x})\rangle=0$,

$$0\in\partial f(\bar{x})+\partial(\bar{\lambda}g)(\bar{x})+B^*\bar{\mu}.$$

## §7 Lagrange 对偶

今将前两节的结果用于问题

$$\min f(x)=\alpha,g(x)\leqslant 0,h(x)=0,x\in D. \tag{1}$$

Lagrange 函数 $L(x,\lambda,\mu)((x,\lambda,\mu)\in D\times Y_+^*\times Z^*)$将起 §6 中 $L(x,y)$的作用. 令 $M=D\cap g^{-1}(Y_-)\cap h^{-1}(0)$. 首先指出如下结论：

1° 利用 2.2.2 易得 $\sup L(x,Y_+^*,Z^*)=f(x)+\delta_M(x)$. 注意到 $\inf(f+\delta_M)(D)=\inf f(M)$，知(1)的对偶问题是

$$\max \inf L(D,\lambda,\mu) = \beta,(\lambda,\mu) \in Y_+^* \times Z^*. \tag{1*}$$

称(1*)为(1)的"Lagrange 对偶". 注意(1*)的约束比较简单,但其目标函数 $\inf L(D,\lambda,\mu)$则未必如此.

2° 极值关系. 与§6(1)相当的等式是 $f(\bar{x})+\delta_M(\bar{x})=\inf L(D,\bar{\lambda},\bar{\mu})$;可验知后者又相当于

$$\begin{cases} f(\bar{x}) = \min L(D,\bar{\lambda},\bar{\mu}); \\ \langle\bar{\lambda},g(\bar{x})\rangle = 0,(\bar{x},\bar{\lambda},\bar{\mu}) \in M \times Y_+^* \times Z^*. \end{cases} \tag{2}$$

结合以上结论与 5.6.1 得出:

**5.7.1 定理** 设$(\bar{x},\bar{\lambda},\bar{\mu})\in M\times Y_+^*\times Z^*$. (i)$(\bar{x},\bar{\lambda},\bar{\mu})$是 $L$ 的鞍点$\Leftrightarrow\bar{x}$ 与$(\bar{\lambda},\bar{\mu})$分别是(1)与(1*)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow$(2)成立. (ii)若$(\bar{x},\bar{\lambda},\bar{\mu})$是 $L$ 的鞍点,$D$ 凸,$f,g,h$ 在 $\bar{x}$ 可微,则

$$\langle L_x(\bar{x},\bar{\lambda},\bar{\mu}),D-\bar{x}\rangle \geqslant 0 = \langle\bar{\lambda},g(\bar{x})\rangle; \tag{3}$$

因此当 $\bar{x}\in D^\circ$时 $\bar{x}$ 是 $K-T$ 点. (iii)若 $\bar{x}$ 是(1)的 $K$-$T$ 点,$\bar{\lambda},\bar{\mu}$ 是 $K-T$ 乘子(依 5.2.1),$L(\cdot,\bar{\lambda},\bar{\mu})$在 $\bar{x}$ 相对于 $D$ 伪凸,则$(\bar{x},\bar{\lambda},\bar{\mu})$是 $L$ 的鞍点.

**证** (i)直接由 5.6.1 推出. 利用 4.1.10(iv)从(2)推出(3). 在(iii)的条件下,(3)必满足;用伪凸假设可从(3)推出(2),故$(\bar{x},\bar{\lambda},\bar{\mu})$是 $L$ 的鞍点. □

更深入的结果基于择一定理. 目前的考虑与导致 5.5.14 的思路颇相类似. 设 $\alpha$ 有限,令 $\tilde{f}(x)=f(x)-\alpha$(与§5 中的 $\tilde{f}(x)$对照!),则问题

$$\exists\, x \in D:\tilde{f}(x) < 0,g(x) \leqslant 0,h(x) = 0 \tag{4}$$

必无解. 若由某个"Motzkin 型定理"(如 5.5.7)得出:

$$\exists\,(\rho,\lambda,\mu) \in (0,\infty) \times Y_+^* \times Z^*:(\rho\tilde{f} + \lambda g + \mu h)(D) \geqslant 0, \tag{5}$$

则可设 $\rho=1$,于是由(5)有 $\alpha\leqslant L(D,\lambda,\mu)$,从而 $\alpha\leqslant\inf L(D,\lambda,\mu)\leqslant\beta$,因此 $\alpha=\beta$ 且$(\lambda,\mu)$是(1*)的解. 基于以上分析,从 5.5.7 得出:

**5.7.2 定理** 设 $\alpha$ 有限，$(f,g,h)$ 满足 5.5.7 之条件 $(H_1)\sim(H_4)$（取 $W=R$），则问题 $(1^*)$ 有解且 $\alpha=\beta$.

下面考虑(1)的特款（在(1)中取 $Z=\{0\}$）：

$$\min f(x)=\alpha, g(x)\leqslant 0, x\in D. \tag{6}$$

令 $L(x,\lambda)=f(x)+\lambda g(x)$ $((x,\lambda)\in D\times Y_+^*)$，则问题(6)的 Lagrange 对偶是（参照 $(1^*)$！）

$$\max \inf L(D;\lambda)=\beta, \lambda\in Y_+^*. \tag{6*}$$

令 $M=D\cap g^{-1}(Y_-)$. 首先由 5.7.1 有

**5.7.3 推论** 设 $(\bar{x},\bar{\lambda})\in M\times Y_+^*$. (i) $(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda)$ 的鞍点 $\Leftrightarrow \bar{x}$ 与 $\bar{\lambda}$ 分别为(6)与 $(6^*)$ 的解且 $\alpha=\beta\neq\pm\infty \Leftrightarrow f(\bar{x})=L(\bar{x},\bar{\lambda})=\min L(D,\bar{\lambda})$. (ii) 若 $(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda)$ 的鞍点，$D$ 凸，$f,g$ 在 $\bar{x}$ 可微，则 $\langle L_x(\bar{x},\bar{\lambda}),D-\bar{x}\rangle\geqslant 0=\langle\bar{\lambda},g(\bar{x})\rangle$，因此当 $\bar{x}\in D^\circ$ 时 $\bar{x}$ 是 $K$-$T$ 点. (iii) 若 $\bar{x}$ 是(6)的 $K$-$T$ 点，$\bar{\lambda}$ 是 $K$-$T$ 乘子，$L(\cdot,\bar{\lambda})$ 在 $\bar{x}$ 相对于 $D$ 伪凸，则 $(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda)$ 的鞍点.

若 $|x|\to\infty$ $(x\in D)\Rightarrow f(x)\to\infty$，则 $L(x,0)$ 亦然. 若 $\hat{x}\in D$, $g(\hat{x})\ll 0$，则 $|\lambda|\to\infty$ $(\lambda\in Y_+^*)\Rightarrow L(\hat{x},\lambda)\to-\infty$（参考 2.2.2(ii)）. 基于此，可用 5.6.2 得出：

**5.7.4 定理** 设 $X,Y$ 自反，$D$ 闭凸，$f$ 凸且 lsc，$g$ 凸且 $*$lsc. 若 $\alpha<\infty$，$|x|\to\infty$ $(x\in D)\Rightarrow f(x)\to\infty$，则(6)有解且 $\alpha=\beta\neq\pm\infty$；若 $\beta>-\infty$，$\exists\ \hat{x}\in D$: $g(\hat{x})\ll 0$，则 $(6^*)$ 有解且 $\alpha=\beta\neq\pm\infty$.

**注** 若 $Z\neq\{0\}$，则对任何 $x\in D$，当 $(\lambda,\mu)\in Y_+^*\times Z^*$，$|\lambda|+|\mu|\to\infty$ 时不能有 $L(x,\lambda,\mu)\to-\infty$，因此不能用 5.6.2 于(1)$(1^*)$ 得出对应于 5.7.4 的结果.

现在考虑对(6)应用择一定理. 设 $\alpha$ 有限，$\alpha$ 是问题(6)的最优值；令 $\mathcal{J}(x)=f(x)-\alpha$，则问题

$$\exists\ x\in D: \langle (R_+\times Y_+^*)\backslash\{0\},(\mathcal{J}(x),g(x))\rangle<0 \tag{7}$$

必无解. 注意形式上(7)恰与 §6(17)一致. 当 $Y^\circ{}_+\neq\varnothing$ 时(7)可改写成：

$$\exists\ x \in D: \bar{f}(x) < 0, g(x) \ll 0. \tag{8}$$

于是只需对$(\bar{f},g)|D$应用 5.6.6(或 5.5.10,若$Y^\circ_+ \neq \varnothing$). 因必要的步骤已见于 5.7.2 与 5.6.7 之证,下面只写出结果.

**5.7.5 定理** 设$\alpha$有限,$\exists\ \hat{x} \in D: \langle Y_+^* \backslash \{0\}, g(\hat{x})\rangle < 0$. 假定以下条件之一满足:(i)$Y^\circ_+ \neq \varnothing$,$(f,g)|D$关于锥$\mathrm{R}_+ \times Y_+$近次类凸,或广义近次类凸且$f(D),g(D)$有界;(ii)$Y_+^*$有弱$^*$紧凸基,$(f,g)|D$关于$\mathrm{R}_+ \times Y_+$近次类凸,或$*$拟凸且$*$ lsc,则$(6^*)$有解且$\alpha = \beta \neq \pm\infty$.

## §8 Rockafellar 对偶

本节所述的对偶理论基于如下的共轭泛函概念.

**5.8.1 定义** 设$f: X \to \overline{\mathrm{R}}, D_f \neq \varnothing$. 称$f^*(u) \triangleq \sup\limits_X [u(x) - f(x)]$$(u \in X^*)$为$f$的共轭泛函. 令$f^{**} = (f^*)^*$,称$\bar{f} \triangleq f^{**}|X$为$f$的闭包.

易见$f^*: X^* \to \overline{\mathrm{R}}$总有定义且是 lsc 凸函数.

**5.8.2 例** 1°设$f(x) = \dfrac{1}{p}|x|^p (x \in X)$,$p > 1$,则不难算出$f^*(u) = \dfrac{1}{2}|u|^q$,$p^{-1} + q^{-1} = 1$.

2°设$X$自反,$A = A^* \in L(X,X^*)$强单调,$f(x) = \dfrac{1}{2}\langle Ax, x\rangle$. $A$必可逆(3.3.8). $\forall\ u \in X^*: f - u$严格凸;$f'(x) - u = 0 \Leftrightarrow x = A^{-1}u$. 由此得出$f^*(u) = \langle u, A^{-1}u\rangle - f(A^{-1}u) = \dfrac{1}{2}\langle u, A^{-1}u\rangle$.

3°设$f(x) = |x| (x \in X)$,则$f^* = \delta_B$,$B = \overline{B}_1(0) \subset X^*$.

4°对$A \subset X$有$\delta_A^* = s_A$,$s_A$依 §4.2(1);若$A$是锥,则$\delta_A^* = \delta_Q$,$Q = -A^*$.

**5.8.3 命题** 设$f,g: X \to \overline{\mathrm{R}}, D_f \neq \varnothing$. (i)"Young 不等式"成立:

$$u(x) \leqslant f(x) + f^*(u), (x,u) \in X \times X^*, \tag{1}$$

且仅当 $u \in \partial f(x)$ 时(1)为等式. (ii) $f=\overline{f} \Leftrightarrow f$ 凸且 lsc $\Rightarrow f^* \not\equiv \infty$. (iii)若 $g$ 凸、lsc 且 $\overline{f} \leqslant g \leqslant f$, 则 $\overline{f}=g$.

**证** (i)的证明是直接的.

(ii)首先设 $f$ 凸且 lsc. 取 $\bar{x}, \bar{t}$ 使 $\bar{t}<f(\bar{x})<\infty$. 对 $(\bar{x}, \bar{t})$ 与 epi$f$ 用分离定理得 $(u,\alpha) \in X^* \times \mathrm{R}$: $u(\bar{x})+\alpha \bar{t}<u(x)+\alpha t$(若 $f(x) \leqslant t$). 这推出 $\alpha>0$, $w=-u/\alpha$ 满足 $f^*(w)<\infty$. 若有 $x_0$ 使 $t_0 \triangleq \overline{f}(x_0)<f(x_0)$, 则有 $v \in X^*$, $\beta, r \in \mathrm{R}$: $v(x_0)+t_0<r<v(x)+\beta(t)$(若 $f(x) \leqslant t$). 必定 $\beta \geqslant 0$. 若 $\beta>0$, 则 $\overline{f}(x_0) \geqslant -\beta^{-1} v(x_0)-f^*(-\beta^{-1} \cdot v)>t_0$, 得出矛盾. 若 $\beta=0$, 则 $v(x_0)<r<v(D_f)$, 于是当 $t \rightarrow \infty$ 时 $\overline{f}(x_0) \geqslant\langle w-t v, x_0\rangle-f^*(w-t v) \geqslant w(x_0)-f^*(w)+t[r-v(x_0)] \rightarrow \infty$, 亦得矛盾. 故必 $f \leqslant \overline{f}$. 其次显然 $\overline{f} \leqslant f$, 因此 $f=\overline{f}$. 反之, 若 $f=\overline{f}$, 则 $f$ 显然凸且 lsc.

(iii)直接由(ii)推出. □

**5.8.4 定理** 设 $X$ 自反, $f: X \rightarrow \mathrm{R}$ 严格凸且 $G$-可微, $\lim\limits_{|x| \rightarrow \infty}|Df(x)|=\infty$, 则 $f^*$ 亦严格凸, $Df^*=(Df)^{-1}$.

**证** $Df: X \rightarrow X^*$ 严格单调且次连续(4.1.5), 故为双射, 且 $F \triangleq (Df)^{-1}: X^* \rightarrow X$ 亦严格单调、次连续(3.3.8). $\forall u, v \in X^*, t \in \mathrm{R}$, 由 5.8.3(i)与 4.1.5(i)有

$$\begin{aligned}\Delta f^*(u, t v) &=[\langle u+t v, F(u+t v)\rangle-f(F(u+t v))] \\ &\quad -[\langle u, F u\rangle-f(F u)] \\ &\leqslant t\langle v, F(u+t v)\rangle .\end{aligned}$$

由此得 $D_+ f^*(u, v) \leqslant\langle v, F u\rangle \leqslant D_- f^*(u, v)$, 因此 $Df^*(u)=F u$, 即 $Df^*=(Df)^{-1}$. 因 $Df^*$ 严格单调, 故 $f^*$ 严格凸. □

Rockafellar 将共轭泛函与扰动函数结合起来, 建立了一个对偶理论, 它蕴涵丰富的结果.

给定 $\varphi: X \times Y \rightarrow \overline{\mathrm{R}}$, 定义所谓 $S$ 泛函如下:

$$S(y)=\inf_x \varphi(x, y), S_*(u)=\inf_\lambda \varphi^*(u, \lambda). \tag{2}$$

其次，由 $\varphi$ 诱导出一个 Lagrange 函数

$$L(x,\lambda) = \inf_y[\varphi(x,y) - \lambda(y)] = -\varphi_y^*(x,\lambda), \tag{3}$$

$\varphi_y^*(x,\cdot)$ 记 $\varphi(x,\cdot)$ 的共轭泛函. 形成一对问题：

$$\min\varphi(x,0) = \alpha, x \in X; \tag{P}$$

$$\max -\varphi^*(0,\lambda) = \beta, \lambda \in Y^*, \tag{$P^*$}$$

称($P^*$)为(P)的 Rockafellar 对偶.

在关于 $\varphi$ 的适当假定下可验知以下结论：

1°$S_*$ 凸；$\varphi$ 凸 $\Rightarrow S$ 凸；$S^*(\lambda)=\varphi^*(0,\lambda)$；

$$S^{**}(0) = -S_*(0) = \beta \leqslant \alpha = S(0). \tag{4}$$

2°$L(x,\cdot)$ 凹且 usc；$\varphi$ 凸[lsc]$\Rightarrow L$ 鞍[半连续].

3°因 $\inf\limits_x L(x,\lambda) = -\varphi^*(0,\lambda)$，故($P^*$)相当于

$$\max\inf_x L(x,\lambda) = \beta, \lambda \in Y^*.$$

其次，若 $\varphi(x,\cdot)$ 凸且 lsc，则由 5.8.3(ii)与(3)有 $\varphi(x,0) = \varphi_y^{**}(x,0) = \sup\limits_\lambda L(x,\lambda)$. 在这种情况下，问题(P)与($P^*$)在 § 6 的意义下互为对偶. 另一方面，若 $X,Y$ 自反，$\varphi$ 凸、lsc，则互换 $\varphi,\varphi^*$ 知(P)是($P^*$)的 Rockafellar 对偶.

**5.8.5 定义** 若 $\partial S(0)\neq\varnothing$，则说问题(P)稳定；若 $\partial S_*(0)\neq\varnothing$，则说问题($P^*$)稳定.

利用稳定性概念，在问题(P)与($P^*$)之间建立了异常和谐的联系. 以下是主要结果.

**5.8.6 定理**(Rockafellar, 1967) 对(P)与($P^*$)有以下结论：(i)$\bar{x}$ 与 $\bar{\lambda}$ 分别为(P)与($P^*$)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow$

$$\varphi(\bar{x},0) + \varphi^*(0,\bar{\lambda}) = 0 \tag{5}$$

$\Leftrightarrow(0,\bar{\lambda})\in\partial\varphi(\bar{x},0)$($\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $L$ 的鞍点，若 $\varphi(x,\cdot)$ 凸、lsc). (ii)若 $\alpha$ 有限，则(P)稳定$\Leftrightarrow$($P^*$)有解且 $\alpha=\beta\Rightarrow\partial S(0)$ 是($P^*$)的解集；若 $X,Y$ 自反，$\varphi$ 凸、lsc，$\beta$ 有限，则($P^*$)稳定$\Leftrightarrow$(P)有解且 $\alpha=\beta\Rightarrow\partial S_*(0)$ 是(P)的解集. (iii)若 $\alpha>-\infty$，$\exists\ \hat{x}$：$\varphi(\hat{x},y)$ 在 $y=0$ 连续，$\varphi$ 凸，则(P)稳定；若 $\beta<\infty$，$\exists\ \dot{\lambda}$：$\varphi^*(u,\dot{\lambda})$ 在 $u=0$ 连续，则($P^*$)稳

定.

**证** (i)的证明是直接的(参照 5.6.1,5.8.3).

(ii)只需证前一半:由(4)与 5.8.3(i)有 $\lambda\in\partial S(0)\Leftrightarrow 0=S(0)+S^*(\lambda)=\alpha+\varphi^*(0,\lambda)\Leftrightarrow\alpha=-\varphi^*(0,\lambda)=\beta$.

(iii)只需证前一半.由 $\varphi(\hat{x},y)$ 在 $y=0$ 连续推出 $S(y)$ 在 $y=0$ 邻近上有界;而 $S$ 凸,$S(0)=\alpha>-\infty$,故 $S(y)$ 必在 $y=0$ 邻近有限,因而在 $y=0$ 连续(4.1.1),于是 $\partial S(0)\neq\varnothing$(4.2.3),即(P)稳定. □

**5.8.7 推论** (i)若 $X,Y$ 自反,$\varphi$ 凸、lsc,则(P)稳定且有解$\Leftrightarrow$(P$^*$)稳定且有解$\Leftrightarrow$(P)(P$^*$)皆稳定且 $\alpha$ 有限$\Leftrightarrow$(P)(P$^*$)皆有解且 $\alpha=\beta\Leftrightarrow L$ 有鞍点.(ii)若 $\bar{x}$ 是(P)的解,$\varphi$ 凸,$\exists\ \hat{x}:\varphi(\hat{x},y)$ 在 $y=0$ 连续,则$\exists\ \bar{\lambda}\in Y^*$ 使条件(5)满足.

**注** 条件(5)与 $K$-$T$ 条件相当(在一定情况下它确与 $K-T$ 条件一致),但它完全不涉及导数.条件"$\exists\ \hat{x}:\varphi(\hat{x},y)$ 在 $y=0$ 连续"可类比于 Slater 条件(在一定情况下它确与 Slater 条件一致),此条件有点强,不少新近研究致力于减弱此条件.

在 Rockafellar 定理的具体应用中,出发点往往不是某个给定的 $\varphi$,而是某个给定的"原问题"(P).关键在于引进适当的"扰动变量"$y$ 并构成"扰动函数"$\varphi(x,y)$,使得问题(P)可表为"$\min\limits_x\varphi(x,0)=\alpha$",从而得到其 Rockafellar 对偶(P$^*$).$\varphi$ 的构成与(P)的目标函数及约束的特性有关,是一个颇带技术性的问题.一些典型的方法在下节考虑.

## §9 Fenchel 对偶

给定 $F:X\to\bar{R}$,$G:Y\to\bar{R}$,$H:Z\to\bar{R}$,$g:X\to Y$,$h:X\to Z$,考虑无约束极小问题

$$\min[F(x)+G(g(x))+H(h(x))]=\alpha,x\in X. \qquad (1)$$

为应用 Rockafellar 理论,构成扰动函数

$$\varphi(x,y,z) = F(x) + G(g(x) + y) + H(h(x) + z). \quad (2)$$

在关于 $F,G,H,g,h$ 的适当假定下可验知以下结论：

1°若 $G,H$ 凸[lsc]，则 $\varphi(x,\cdot,\cdot)$ 凸[lsc]；若 $F,G,H,g$ 凸，$G$ 单调增（依 $Y_+$ 导出的序 $\leqslant$），$h(x)=Bx-b$，则 $\varphi$ 凸.

2°令 $\Gamma(\cdot,\lambda,\mu)=F+\lambda g+\mu h$，则经直接计算得出：

$$\varphi^*(u,\lambda,\mu) = \Gamma_x^*(u,\lambda,\mu) + G^*(\lambda) + H^*(\mu) \quad (3)$$

($(u,\lambda,\mu)\in X^*\times Y^*\times Z^*$). 于是(1)的 Rockafellar 对偶为

$$\max[-\Gamma_x^*(0,\lambda,\mu) - G^*(\lambda) - H^*(\mu)] = \beta, \lambda\in Y^*, \mu\in Z^*. \quad (1^*)$$

3°结合(2)(3)有

$$\begin{aligned}\varphi(x,0,0) + \varphi^*(0,\lambda,\mu) = &[\Gamma(x,\lambda,\mu) + \Gamma_x^*(0,\lambda,\mu) - \langle 0,x\rangle] \\ &+ [G(g(x)) + G^*(\lambda) - \langle\lambda,g(x)\rangle] \\ &+ [H(h(x)) + H^*(\mu) - \langle\mu,h(x)\rangle].\end{aligned}$$

上式右端三项皆非负(5.8.3(i))，于是对应于 §8(5)的条件“$\varphi(\bar{x},0,0)+\varphi^*(0,\bar{\lambda},\bar{\mu})=0$”可写成

$$0\in\partial_x\Gamma(\bar{x},\bar{\lambda},\bar{\mu}), \bar{\lambda}\in\partial G(g(\bar{x})), \bar{\mu}\in\partial H(h(\bar{x})). \quad (4)$$

其中 $0\in\partial_x\Gamma(\bar{x},\bar{\lambda},\bar{\mu})\Leftrightarrow\Gamma(\bar{x},\bar{\lambda},\bar{\mu})=\min\limits_x\Gamma(x,\bar{\lambda},\bar{\mu})\Leftrightarrow\Gamma(\bar{x},\bar{\lambda},\bar{\mu})+\Gamma_x^*(0,\bar{\lambda},\bar{\mu})=0$.

4°若令 $\Lambda(x,\lambda,\mu)=-\varphi_{(y,z)}^*(x,\lambda,\mu)$（参考 §8(3)），则

$$\Lambda(x,\lambda,\mu) = \begin{cases}\Gamma(x,\lambda,\mu) - G^*(\lambda) - H^*(\mu), & x\in D_F; \\ \infty, & \text{否则}.\end{cases} \quad (5)$$

5°若 $\hat{x}\in D_F$，$G,H$ 分别在 $g(\hat{x})$ 与 $h(\hat{x})$ 连续，则 $\varphi(\hat{x},\cdot,\cdot)$ 在(0,0)连续；若 $G^*(\hat{\lambda})$ 与 $H^*(\hat{\mu})$ 有限，$\Gamma_x^*(u,\hat{\lambda},\hat{\mu})$ 在 $u=0$ 连续，则 $\varphi^*(u,\hat{\lambda},\hat{\mu})$ 在 $u=0$ 连续.

结合以上结论与 5.8.6 得出：

**5.9.1 定理** (i) $\bar{x}$ 与 $(\bar{\lambda},\bar{\mu})$ 分别为(1)与(1*)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow$(4)成立($\Leftrightarrow(\bar{x},\bar{\lambda},\bar{\mu})$ 是 $\Lambda$ 的鞍点，若 $G,H$ 凸且 lsc). (ii)若 $\alpha$ 有限，则(1)稳定$\Leftrightarrow$(1*)有解且 $\alpha=\beta$；若 $X,Y,Z$ 自反，$F,G,H,g$

凸，$G$ 单调增，$h(x)=Bx-b$，$\varphi$ lsc，$\beta$ 有限，则$(1^*)$稳定$\Leftrightarrow$(1)有解且 $\alpha=\beta$. (iii)若 $\alpha>-\infty$，$F,G,H,g$ 凸，$G$ 单调增，$h(x)=Bx-b$，$\exists\ \hat{x}\in D_F$：$G,H$ 分别在 $g(\hat{x})$ 与 $h(\hat{x})$ 连续，则(1)稳定；若 $\beta<\infty$，$\exists\ \hat{\lambda},\hat{\mu}$：$G^*(\hat{\lambda})$ 与 $H^*(\hat{\mu})$ 有限且 $\Gamma_x^*(u,\hat{\lambda},\hat{\mu})$ 在 $u=0$ 连续，则$(1^*)$稳定.

问题(1)有很大的一般性，其具体应用取决于 $F,G,H$ 的特殊选择. 若取 $F=f+\delta_D$，$G=\delta_{Y_-}$，$H=\delta_{\{0\}}(0\in Z)$，则问题(1)成为：

$$\min f(x)=\alpha,g(x)\leqslant 0,h(x)=0,x\in D. \tag{6}$$

注意 $G^*=\delta_{Y_+^*}$(5.8,2)，$H^*=0$. 令 $L(\cdot,\lambda,\mu)=f+\lambda g+\mu h$，则 $\Gamma(\cdot,\lambda,\mu)=L(\cdot,\lambda,\mu)+\delta_D$. (2)(3)(5)依次成为：

$$\varphi(x,y,z)=\begin{cases}f(x),x\in D,g(x)\leqslant -y,h(x)=-z,\\ \infty,\quad 否则；\end{cases} \tag{7}$$

$$\varphi^*(u,\lambda,\mu)=\begin{cases}\sup\limits_{x\in D}[u(x)-L(x,\lambda,\mu)],\lambda\in Y_+^*,\\ \infty,\qquad\qquad 否则；\end{cases} \tag{8}$$

$$\Lambda(x,\lambda,\mu)=\begin{cases}L(x,\lambda,\mu),x\in D,\lambda\in Y_+^*,\\ -\infty,\qquad x\in D,\lambda\,\overline{\in}\,Y_+^*,\\ \infty,\qquad x\,\overline{\in}\,D.\end{cases} \tag{9}$$

由(8)知(6)的对偶是：

$$\max\inf L(D,\lambda,\mu)=\beta,(\lambda,\mu)\in Y_+^*\times Z^*, \tag{6$^*$}$$

这正是(6)的 Lagrange 对偶(见§7$(1^*)$)！令 $M=D\cap g^{-1}(Y_-)\cap h^{-1}(0)$，不难验知条件(4)现在可写成：

$$\begin{cases}f(\bar{x})=\min L(D,\bar{\lambda},\bar{\mu});\\ \langle\bar{\lambda},g(\bar{x})\rangle=0,(\bar{x},\bar{\lambda},\bar{\mu})\in M\times Y_+^*\times Z^*,\end{cases} \tag{10}$$

这正是§7 中的条件(2)！注意到 $G=\delta_{Y_-}$ 与 $H=\delta_{\{0\}}$ 凸且 lsc，$G$ 单调增，从 5.9.1 得出：

**5.9.2 推论** (i)$\bar{x}$ 与$(\bar{\lambda},\bar{\mu})$分别为(6)与$(6^*)$的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow$(10)成立$\Leftrightarrow(\bar{x},\bar{\lambda},\bar{\mu})$是 $L|D\times Y_+^*\times Z^*$ 的鞍点. (ii)若 $\alpha$ 有限，则(6)稳定$\Leftrightarrow(6^*)$有解且 $\alpha=\beta$；若 $X,Y,Z$ 自反，$f$ 凸且 lsc，$g$

凸且 $*$ lsc，$D$ 闭凸，$h(x)=Bx-b$，$\beta$ 有限，则(6*)稳定⇔(6)有解且 $\alpha=\beta$. (iii) 若 $\beta<\infty$，$\exists\ (\hat{\lambda},\hat{\mu})\in Y_+^*\times Z^*$：$u\mapsto \sup\limits_{x\in D}[u(x)-L(x,\hat{\lambda},\hat{\mu})]$ 在 $u=0$ 连续，则(6*)稳定.

注意 5.9.2 之结论(i)与 5.7.1(i)完全一致.

若取 $Z=\{0\}$，则以上讨论略有简化：(1)成为

$$\min[F(x)+G(g(x))]=\alpha, x\in X; \tag{11}$$

相应地，$\varphi(x,y)=F(x)+G(g(x)+y)$，$\Gamma(\cdot,\lambda)=F+\lambda g$，$\varphi^*(u,\lambda)=\Gamma_x^*(u,\lambda)+G^*(\lambda)$，于是(11)的对偶问题是：

$$\max[-\Gamma_x^*(0,\lambda)-G^*(\lambda)]=\beta, \lambda\in Y^*. \tag{11*}$$

条件(4)"缩短"为

$$0\in\partial_x\Gamma(\bar{x},\bar{\lambda}), \bar{\lambda}\in\partial G(g(\bar{x})), \tag{12}$$

其中 $0\in\partial_x\Gamma(\bar{x},\bar{\lambda})\Leftrightarrow\Gamma(\bar{x},\bar{\lambda})=\min\limits_x\Gamma(x,\bar{\lambda})\Leftrightarrow\Gamma(\bar{x},\bar{\lambda})+\Gamma_x^*(0,\bar{\lambda})=0$. 其次，$\Lambda(x,\lambda)=\Gamma(x,\lambda)-G^*(\lambda)\ (x\in D_F)$. 于是有

**5.9.3 推论** (i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为(11)与(11*)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow$(12)成立($\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $\Lambda(x,\lambda)$ 的鞍点，若 $G$ 凸且 lsc). (ii) 若 $\alpha$ 有限，则(11)稳定⇔(11*)有解且 $\alpha=\beta$；若 $X,Y$ 自反，$F,G,g$ 凸，$G$ 单调增，$\varphi$ lsc，$\beta$ 有限，则(11*)稳定⇔(11)有解且 $\alpha=\beta$. (iii) 若 $\alpha>-\infty$，$F,G,g$ 凸，$G$ 单调增，$\exists\ \hat{x}\in D_F$：$G$ 在 $g(\hat{x})$ 处连续，则(11)稳定；若 $\beta<\infty$，$\exists\ \hat{x}$：$G^*(\bar{\lambda})$ 有限，$\Gamma_x^*(u,\hat{\lambda})$ 在 $u=0$ 连续，则(11*)稳定.

将 5.9.3 用于问题

$$\min f(x)=\alpha, g(x)\leqslant 0, x\in D \tag{13}$$

与

$$\max\inf L(D,\lambda)=\beta, \lambda\in Y_+^*, \tag{13*}$$

其中 $L(\cdot,\lambda)=f+\lambda g$，得出与 5.9.2 对应的以下结论：

**5.9.4 推论** (i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为(13)与(13*)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow f(\bar{x})=L(\bar{x},\bar{\lambda})=\min L(D,\bar{\lambda})$，$(\bar{x},\bar{\lambda})\in(D\cap g^{-1}(Y_-))\times Y_+^*\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $L|D\times Y_+^*$ 的鞍点. (ii) 若 $\alpha$ 有限，则(13)稳定⇔(13*)有解且 $\alpha=\beta$；若 $X,Y$ 自反，$f$ 凸且 lsc，$g$ 凸且 $*$ lsc，$D$ 闭凸，

$\beta$ 有限，则(13*)稳定$\Leftrightarrow$(13)有解且 $\alpha=\beta$. (iii)若 $\alpha>-\infty, f, g, D$ 凸，$\exists\ \hat{x}\in D: g(\hat{x})\ll 0$，则(13)稳定；若 $\beta<\infty$，$\exists\ \bar{\lambda}\in Y_+^*: u\mapsto \sup\limits_{x\in D}[u(x)-L(x,\hat{\lambda})]$ 在 $u=0$ 连续，则(13*)稳定.

今在(11)中取 $g(x)=Ax-a$，即考虑问题

$$\min[F(x)+G(Ax-a)]=\alpha, x\in X. \tag{14}$$

令 $\Gamma(x,\lambda)=F(x)+\langle\lambda,Ax-a\rangle$，则 $\Gamma_x^*(u,\lambda)=F^*(u-A^*\lambda)+\lambda(a)$. 以此代入(11*)得出(14)的对偶问题：

$$\max[-F^*(-A^*\lambda)-G^*(\lambda)-\lambda(a)]=\beta, \lambda\in Y^*. \tag{14*}$$

通常称(14*)为(14)的"Fenchel 对偶". 在广义的意义上，我们称(1*)为(1)的 Fenchel 对偶. 直接验知 $\partial_x\Gamma(x,\lambda)=\partial F(x)+A^*\lambda$ (参考 4.2.5)，于是由 5.9.3 得出：

**5.9.5 推论** (i)$\bar{x}$ 与 $\bar{\lambda}$ 分别为(14)与(14*)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow -A^*\bar{\lambda}\in\partial F(\bar{x})$ 且 $\bar{\lambda}\in\partial G(A\bar{x}-a)$ ($\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $\Gamma(x,\lambda)-G^*(\lambda)$ 的鞍点，若 $G$ 凸且 lsc). (ii)若 $\alpha$ 有限，则(14)稳定$\Leftrightarrow$(14*)有解且 $\alpha=\beta$；若 $X, Y$ 自反，$F, G$ 凸且 lsc，$\beta$ 有限，则(14*)稳定$\Leftrightarrow$(14)有解且 $\alpha=\beta$. (iii)若 $\alpha>-\infty$，$F, G$ 凸，$\exists\ \hat{x}\in D_F: G$ 在点 $A\hat{x}-a$ 连续，则(14)稳定；若 $\beta<\infty$，$\exists\ \hat{\lambda}: G^*(\hat{\lambda})$ 有限且 $F^*(u)$ 在 $u=-A^*\hat{\lambda}$ 连续，则(14*)稳定.

若在(14)中 $F\in X^*$，则 $F^*=\delta_{\{F\}}$，代入(14*)得

$$\max[-G^*(\lambda)-\lambda(a)]=\beta, A^*\lambda+F=0. \tag{14)*}$$

结合 5.8.4 与 5.9.5 可建立：

**5.9.6 定理** 设 $F\in X^*$，$G: Y\to\mathbb{R}$ 严格凸且 $G$-可微，则(i)$\bar{x}$ 与 $\bar{\lambda}$ 分别为(14)与(14)* 的解且 $\alpha=\beta\Leftrightarrow A^*\bar{\lambda}+F=0$ 且 $\bar{\lambda}=DG(A\bar{x}-a)\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $F(x)+\lambda(Ax-a)-G^*(\lambda)$ 的鞍点；(ii)若 $X$ 自反，$DG$ 强制，$|Ax|\geqslant C|x|$ $(0<C=\mathrm{const})$，则(14)与(14)*分别有唯一解且 $\alpha=\beta$.

**证** 只需证(ii). 所给条件推出 $x\mapsto A^*DG(Ax-a)$ 严格单调

且强制，因此有唯一 $\bar{x}\in X$，使 $\bar{\lambda}\triangleq DG(A\bar{x}-a)$ 满足 $A^*\bar{\lambda}+F=0$ (3.3.8)。由结论(i)，$\bar{x}$ 与 $\bar{\lambda}$ 分别为(14)与(14)$^*$的解且 $\alpha=\beta$。因 $G$ 与 $G^*$ 皆严格凸(5.8.4)，从而(14)与(14)$^*$的目标函数分别为严格凸与严格凹，故得唯一性结论。 □

若在(14)中取 $A=\mathrm{id}$，$X=Y$，$a=0$，则得所谓"和极小问题"：

$$\min[F(x)+G(x)]=\alpha, x\in X. \tag{15}$$

(15)的 Fenchel 对偶是(参照(14$^*$))：

$$\max[-F^*(-u)-G^*(u)]=\beta, u\in X^*. \tag{15$^*$}$$

直接由 5.9.5 有

**5.9.7 推论** (i) $\bar{x}$ 与 $\bar{u}$ 分别为(15)与(15$^*$)的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow\bar{u}\in(-\partial F(\bar{x}))\cap\partial G(\bar{x})$ ($\Leftrightarrow(\bar{x},\bar{u})$ 是 $F(x)-G^*(u)-u(x)$ 的鞍点，若 $G$ 凸且 lsc)。(ii)若 $\alpha$ 有限，则(15)稳定$\Leftrightarrow$(15$^*$)有解且 $\alpha=\beta$；若 $X$ 自反，$F,G$ 凸且 lsc，$\beta$ 有限，则(15$^*$)稳定$\Leftrightarrow$(15)有解且 $\alpha=\beta$。(iii)若 $\alpha>-\infty$，$F,G$ 凸，$G$ 在某点 $\hat{x}\in D_F$ 连续，则(15)稳定；若 $\beta<\infty$，$\exists\,\hat{u}$：$G^*(\hat{u})$ 有限，$F^*(u)$ 在 $u=-\hat{u}$ 连续，且(15$^*$)稳定。

## §10 线性与二次最优化

线性最优化问题最早得到研究，且已形成很完整的理论，至少在有限维情况下是如此；它的一些方法为非线性最优化理论所借鉴。现在我们感兴趣的问题是：一般的"非线性理论"用于线性问题能导致哪些结果？

以下设 $X,X^*,Y,Y^*$ 中分别由 $X_+,X_+^*,Y_+,Y_+^*$ 导入的序皆记作$\leqslant$。给定 $f\in X^*$，考虑线性最优化问题

$$\min f(x)=\alpha, Ax\leqslant a, x\geqslant 0. \tag{1}$$

今考虑对(1)应用 5.9.4。依 5.9.4 的记号有 $g(x)=Ax-a$，$L(x,\lambda)=f(x)+\lambda(Ax-a)$，$D=X_+$，

$$\inf L(D,\lambda)=\inf(A^*\lambda+f,X_+)-\lambda(a)$$

$$=\begin{cases}-\lambda(a), & A^*\lambda+f\geqslant 0;\\ -\infty, & \text{否则},\end{cases}$$

因此(1)的对偶问题是

$$\max -\lambda(a)=\beta, A^*\lambda+f\geqslant 0, \lambda\geqslant 0. \tag{1*}$$

注意(1*)是与(1)同类型的线性最优化问题. 分别以 $M,N$ 记(1)与(1*)的可行集, 则 $M\neq\varnothing\Rightarrow\alpha<\infty$, $N\neq\varnothing\Rightarrow\beta>-\infty$. 条件"$f(\bar{x})=\min L(D,\bar{\lambda}), \langle\bar{\lambda},g(\bar{x})\rangle=0, (\bar{x},\bar{\lambda})\in M\times Y_+^*$"可表成具对称性的以下形式:

$$\begin{cases}\langle\bar{\lambda},A\bar{x}-a\rangle=\langle A^*\bar{\lambda}+f,\bar{x}\rangle=0;\\ \bar{x}\geqslant 0,\bar{\lambda}\geqslant 0,A\bar{x}\leqslant a,A^*\bar{\lambda}+f\geqslant 0.\end{cases} \tag{2}$$

因

$$\varphi(x,y)=\begin{cases}f(x), Ax+y\leqslant a, x\geqslant 0;\\ \infty, \quad \text{否则},\end{cases}$$

不难依 §8(2)算得

$$S(y)=\inf\{f(x):Ax\leqslant a-y, x\geqslant 0\}\triangleq\alpha(a-y);$$

$$S_*(u)=-\sup\{-\lambda(a):A^*\lambda+f-u\geqslant 0,\lambda\geqslant 0\}$$

$$\triangleq-\beta(f-u).$$

据此得 $\partial S(0)=-\partial\alpha(a)$, $\partial S_*(0)=\partial\beta(f)$. 综合以上结论并用 5.9.4, 5.8.6 得出:

**5.10.1 定理** (i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为(1)与(1*)的解且 $\alpha=\beta\Leftrightarrow$(2)成立$\Leftrightarrow(\bar{x},\bar{\lambda})$是 $L(x,\lambda)$在 $X_+\times Y_+^*$ 上的鞍点. (ii)若 $M,N\neq\varnothing$, 则(1)稳定$\Leftrightarrow$(1*)有解且 $\alpha=\beta\Rightarrow-\partial\alpha(a)$是(1*)的解集, 而当 $X,Y$ 自反时(1*)稳定$\Leftrightarrow$(1)有解且 $\alpha=\beta\Rightarrow\partial\beta(f)$是(1)的解集. (iii)若 $N\neq\varnothing$, $\exists\ \hat{x}\geqslant 0: A\hat{x}\ll a$, 则(1)稳定; 若 $M\neq\varnothing$, $\exists\ \hat{\lambda}\geqslant 0: A^*\hat{\lambda}+f\gg 0$, 则(1*)稳定.

注意条件(2)正是线性规划中有名的"松弛互补条件". 5.10.1 不失为一个漂亮的结果, 但它并未达到线性规划中相应结果的完满程度, 因为有以下更强的有限维结果:

**5.10.2 定理** 设 $X=\mathrm{R}^n, Y=\mathrm{R}^m, X_+=\mathrm{R}_+^n, Y_+=\mathrm{R}_+^m$. 若 $M,N$

$\neq\varnothing$，则问题(1)与(1*)皆有解、稳定且 $\alpha=\beta$；若 $M=\varnothing$ 或 $N=\varnothing$，则(1)与(1*)皆无解.

以上结果的证明基于某些精细的线性代数方法，这些方法无法推广于无限维情况.

现在转而考虑二次最优化问题

$$\min\left[\frac{1}{2}\langle Qx,x\rangle+w(x)\right]=\alpha, Ax\leqslant a, \tag{3}$$

其中 $Q=Q^*\in L(X,X^*)$，$w\in X^*$，假定 $X,Y$ 自反. 令 $f(x)=2^{-1}\langle Qx,x\rangle+w(x)$，$L(x,\lambda)=f(x)+\langle\lambda,Ax-a\rangle$. 设 $Q$ 强单调，于是 $Q$ 可逆，$f$ 严格凸，从而 $L(\cdot,\lambda)$ 有唯一极小点 $x$，它决定于 $L_x(x,\lambda)=Qx+A^*\lambda+w=0$. 因此

$$\begin{aligned}\min_x L(x,\lambda)&=L(Q^{-1}(-A^*\lambda-w),\lambda)\\&=-\frac{1}{2}\langle\lambda,P\lambda\rangle-\langle\lambda,b\rangle-c,\end{aligned}$$

其中 $P=AQ^{-1}A^*=P^*\in L(Y^*,Y)$，$b=AQ^{-1}w+a$，$c=\frac{1}{2}\langle w,Q^{-1}w\rangle$. 于是(3)的 Lagrange 对偶为

$$\max\left[-\frac{1}{2}\langle\lambda,P\lambda\rangle-\langle\lambda,b\rangle-c\right]=\beta, \lambda\geqslant 0. \tag{3*}$$

注意(3*)的目标函数与(3)属同一类型，而约束则较简单. 以 $M$ 记(3)的可行集，则 $M\neq\varnothing\Leftrightarrow\alpha$ 有限. 另一方面，恒有 $\beta>-\infty$. 关于(3)的 $K$-$T$ 条件是：

$$\begin{cases}Q\bar{x}+A^*\bar{\lambda}+w=\langle\bar{\lambda},A\bar{x}-a\rangle=0;\\ A\bar{x}\leqslant a,\bar{\lambda}\geqslant 0.\end{cases} \tag{4}$$

由 5.9.4(i)知(4)$\Leftrightarrow(\bar{x},\bar{\lambda})$为 $L$ 的鞍点.

类似于问题(1)，对于问题(3)的 $S$ 泛函亦有 $\partial S(0)=-\partial\alpha(a)$，$\partial S_*(0)=\partial\beta(w)$. 结合以上结论与 5.7.3、5.9.4 得出：

**5.10.3 定理** 设 $X,Y$ 自反，$Q=Q^*\in L(X,X^*)$ 强单调，$w\in X^*$. (i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为(3)与(3*)的解且 $\alpha=\beta\Leftrightarrow$(4)成立$\Leftrightarrow(\bar{x},\bar{\lambda})$是 $2^{-1}\langle Qx,x\rangle+w(x)+\langle\lambda,Ax-a\rangle$ 在 $X\times Y_+^*$ 上的鞍点. (ii)若∃ $\hat{x}\in$

$X: A\hat{x} \ll a$,则(3)稳定,(3*)有非空解集 $-\partial\alpha(a)$;若 $M \neq \varnothing$,则(3*)稳定,(3)有非空解集 $\partial\beta(w)$.

定理中的"强单调"条件无疑过强了.在某些特殊情况下,可以减弱这一条件.

**5.10.4 定理** 设 $Y=\mathbb{R}^m$.(i)若 $\bar{x}$ 是(3)的局部最优解,$R(A)=\mathbb{R}^m$,则有 $\bar{\lambda}\in\mathbb{R}^m$ 使 $(\bar{x},\bar{\lambda})$ 满足(4).(ii)若 $X=\mathbb{R}^n$,$A=(A_1,\cdots,A_m)$,$a=(a_1,\cdots,a_m)$,$(\bar{x},\bar{\lambda})$ 满足(4),当 $z\in\mathbb{R}^n\setminus\{0\}$ 满足"$A_i\bar{x}=a_i\Rightarrow\bar{\lambda}_iA_iz=0$"时 $(Qz,z)>0$,则 $\bar{x}$ 是(3)的严格的局部最优解.

**证** (i)由 5.2.6 推出;(ii)由 5.3.1 推出. □

## §11 最佳逼近问题

设 $D\subset X$ 非空凸.考虑"最佳逼近问题"

$$\min|y-x|=d_D(x),y\in D, \tag{1}$$

$x\in X$ 是给定的.这种问题在不同的具体形式下出现于多个领域.(1)的目标函数似乎很简单,但即使对不很复杂的 $D$,问题也可能很困难,深入的讨论非专著莫能尽其详.此处仅从应用本章方法的角度给出某些基本结论.

以 $Px$ 记(1)的解集.显然 $Px$ 有界、凸,当 $D$ 闭时 $Px$ 亦为闭集.若 $X$ 自反(或 $\dim \operatorname{span} D<\infty$),$D$ 闭,则 $Px\neq\varnothing$(4.1.10);若再假定 $X$ 严格凸,则 $P:X\to D$ 为单值映射.以下记 $B=\bar{B}_1(0)\subset X$,$B^*=\bar{B}_1(0)\subset X^*$.

**5.11.1 定理** 设 $x\in D^c$,$y\in D$.(i)$y\in Px\Leftrightarrow\exists\ u\in\partial B^*\cap(D-y)^*:u(y-x)=|y-x|\Leftrightarrow\exists\ v\in(D-y)^*\setminus\{0\}:v(y-x)=|v||y-x|$.(ii)若 $D$ 是子空间,则 $y\in Px\Leftrightarrow\exists\ u\in\partial B^*\cap D^\perp;u(x)=|y-x|\Leftrightarrow\exists\ v\in D^\perp\setminus\{0\}:v(x)=|v||y-x|$.(iii)若 $X^*$ 严格凸,则 $y\in Px\Leftrightarrow(D-y,y-x)_+\geqslant0$.

**证** (i)$y\in Px\Leftrightarrow\partial_y|y-x|\cap(D-y)^*\neq\varnothing\Leftrightarrow\exists\ u\in\partial B^*\cap(D-y)^*:u(y-x)=|y-x|$(4.2.2,4.2.6),这得出(i).由(i)推出

(ii).

(iii)由 4.1.10(iv)及 3.2.5(ii), $y\in Px$ 等价于

$$0\leqslant|y-x|\langle D,|y-x|,D-y\rangle=\langle\mathscr{D}(y-x),D-y\rangle$$
$$=(D-y,y-x)_+.$$
□

5.11.1(i)可表述为: $y$ 是(1)的解当且仅当

$$\max\langle u,y-x\rangle=\beta,u\in\partial B^*\cap(D-y)^* \quad (2)$$

有解且 $\beta=d_D(x)$; 当 $D$ 是子空间时(2)成为

$$\max u(x)=\beta,u\in\partial B^*\cap D^{\perp}. \quad (3)$$

(2)(或(3))是(1)的某种对偶问题. 对偶问题的约束可能更复杂, 但有线性目标函数.

**5.11.2 推论** 若 $X$ 是 Hilbert 空间, $D$ 闭, 则 $P:X\to D$ 为单调算子且 $\mathrm{Lip}P\leqslant 1$.

**证** $\forall x,y\in X$, 由 5.11.1(iii)有 $(Py-Px,Px-x)\geqslant 0$, $(Px-Py,Py-y)\geqslant 0$. 组合此两不等式得

$$0\leqslant(Px-Py,x-y)-|Px-Py|^2\leqslant(Px-Py,x-y),$$

于是 $|Px-Py|\leqslant|x-y|$. 由此得出所要结论. □

以下这个颇为别致的结果是 Moreau 一个 Hilbert 空间定理(1962)的推广, 它揭示了算子 $P$ 的"投影"性质.

**5.11.3 定理** 设 $X$ 自反, $X$ 与 $X^*$ 严格凸, $D$ 为闭凸锥, $D^+=\{y\in X:(D,y)_+\leqslant 0\}$, 则每个 $x\in X$ 有唯一分解 $x=y+x^+$, 使 $x^+\in D^+$, $(y,x^+)_+=0$; 必定 $y=Px$.

**证** 若 $x=y+x^+$ 是定理所述分解, 则 $(D-y,y-x)_+=(D,-x^+)_+\geqslant 0$, 因此 $y=Px$(5.11.1(iii)), 可见分解是唯一的. 其次, $\forall x\in X$, 令 $y=Px$, $x^+=x-y$, 则 $x=y+x^+$,

$$(D,x^+)_+=-(D+y-y,y-x)_+\leqslant 0;$$

$$0\leqslant(0-y,y-x)_+=(y,x^+)_+=(2y-y,x-y)_+\leqslant 0,$$

这表明 $x^+\in D^+$, $(y,x^+)_+=0$. □

**注** 若 $X$ 是 Hilbert 空间, $D$ 是闭子空间, 则 5.11.3 中的分解正是通常的正交分解.

下面设 $D$ 闭凸并固定 $x\in D^c$. 今对(1)运用 Rockafellar 理论，为此，作扰动函数

$$\varphi(y,z)=|y-x-z|+\delta_D(y),y,z\in X \tag{4}$$

显然 $\varphi$ 凸、lsc，$S(z)=\inf_y\varphi(y,z)=d_D(x+z)$，$\partial S(0)=\partial d_D(x)\neq\varnothing$.
其次，直接计算得出

$$\varphi^*(u,v)=s_D(u+v)-v(x)+\delta_{B^*}(v),u,v\in X^*; \tag{5}$$

$$L(y,v)=-\varphi_z^*(y,v)=v(x-y),(y,v)\in D\times B^*; \tag{6}$$

$$S_*(u)=\inf_v\varphi^*(u,v)=\inf_{|v|\leqslant 1}(\sup\langle u+v,D\rangle-v(x)),u\in X^*, \tag{7}$$

(5)中 $s_D$ 依 §4.2(1). 由(5)知(1)的 Rockafellar 对偶为：

$$\max[v(x)-s_D(v)]=\beta,v\in B^*. \tag{1*}$$

由 $\varphi(y,0)+\varphi^*(0,v)=|y-x|+s_D(v)-v(x)+\delta_D(y)+\delta_{B^*}(v)$ 得出：条件 $\varphi(y,0)+\varphi^*(0,v)=0$(参考 §8(5))等价于

$$\langle v,x-y\rangle=|x-y|,y\in D,v\in\partial B^*\cap(y-D)^*. \tag{8}$$

(8)意味着当 $\beta=|x-y|$ 时 $u=-v$ 是问题(2)的解. 综合以上结论与 5.8.6 得出：

**5.11.4 定理** 设 $D$ 闭凸，$x\in D^c$. (i) $\bar{y}$ 与 $\bar{v}$ 分别为(1)与(1*)的解且 $\beta=d_D(x)\Leftrightarrow$ 当 $y=\bar{y}\in D,\beta=|x-\bar{y}|$ 时 $u=-\bar{v}$ 是(2)的解 $\Leftrightarrow(\bar{y},\bar{v})$ 是 $L(y,v)=\langle v,x-y\rangle$ 在 $D\times B^*$ 上的鞍点. (ii)(1)稳定，$\partial d_D(x)$ 是(1*)的非空解集且 $\beta=d_D(x)$；若 $X$ 自反，则(1*)稳定且 $\partial S_*(0)$ 是(1)的非空解集，$S_*$ 依(7).

下面设 $D=\mathrm{span}\{e_1,e_2,\cdots,e_n\}$ 是 $X$ 的 $n$ 维子空间，$x\in D^c$，$y\in D$. 当 $X$ 是 Hilbert 空间时，$Px$ 是 $x$ 在 $D$ 上的正交投影，它可由熟知的公式表出. 若 $X$ 是凸性"很差"的空间(如 $C(\Omega)$)，则 $P$ 没有单值性且可能有很复杂的性态. 在这种情况下，我们希望将 5.11.1 的条件具体化，这可利用"极端点"概念来实现. 任给 $M\subset X$，称 $x\in M$ 为 $M$ 的极端点，若它不是 $M$ 中任何线段的内点.

**5.11.5 引理** $y\in Px\Leftrightarrow$ 存在 $B^*$ 的 $m(\leqslant n+1)$ 个线性无关的

极端点 $u_j$ 与 $\lambda_j>0$，使 $\Sigma\lambda_j=1$，$u\triangleq\Sigma\lambda_j u_j\in\partial B^*\cap D^\perp$ 且 $u_j(y-x)=|y-x|(1\leqslant j\leqslant m)$.

**证** 设 $y\in Px$，$Y=\text{span}\{x,D\}$. 由 5.11.1，有 $u\in\partial B^*\cap D^\perp$：$u(y-x)=|y-x|$. 由 Krein-Milman 定理[75;1.7.3]，$u|Y$ 可表为 $Y^*\cap B^*$ 的 $m(\leqslant n+1)$ 个线性无关极端点的凸组合：$u=\Sigma\lambda_j u_j$，$\lambda_j>0$；不妨设 $u_j$ 是 $B^*$ 的极端点且 $|u_j|=1$. 由 $\sum\lambda_j u_j(y-x)=|y-x|$ 与 $u_j(y-x)\leqslant|y-x|$ 推出 $u_j(y-x)=|y-x|(1\leqslant j\leqslant m)$. 逆命题直接由 5.11.1 得出. □

若对 $B^*$ 中任何几个线性无关的极端点 $u_i$ 与 $a_i\in\mathbb{R}$，方程组 $u_i(z)=a_i(1\leqslant i\leqslant n)$ 在 $D$ 中有唯一解 $z$，则说 $D$ 有"插值性质"，缩记作(IP).

**5.11.6 例** 设 $X=C(J)$，$e_i(t)=t^{i-1}(1\leqslant i\leqslant n)$，则 $D=\text{span}\{e_1,e_2,\cdots,e_n\}$ 是 $J$ 上次数$<n$ 的多项式之全体. $B^*$ 的极端点形如 $\pm\delta_t(t\in J)$，$\delta_t$ 是集中于点 $t$ 的 Dirac 测度(参考[202; Ⅱ]). 若 $0\leqslant t_1<t_2<\cdots<t_n\leqslant 1$，$a_i\in\mathbb{R}$，则由熟知的 Lagrange 插值公式知有唯一 $z\in D$，使 $z(t_i)=a_i(1\leqslant t\leqslant n)$. 这正表明 $D$ 有(IP).

**5.11.7 引理** 若 $D$ 有(IP)，则 $P:X\to D$ 是单值的.

**证** 设 $y_k\in Px$，$k=1,2$，则 $y=2^{-1}(y_1+y_2)\in Px$，设 $u_j,\lambda_j(1\leqslant j\leqslant m)$ 如 5.11.5. 若 $m\leqslant n$，则 $\{u_j\}$ 可补足为 $B^*$ 的几个线性无关极端点 $u_1,\cdots,u_n$. 由(IP)，有唯一 $z\in D$，使 $u_i(z)=1(1\leqslant i\leqslant n)$. 这推出 $0=\Sigma_1^m\lambda_j u_j(z)=\Sigma_1^m\lambda_j$，得出矛盾. 因此必 $m=n+1$. 由

$$|y-x|=u_j(y-x)=\frac{1}{2}\sum_k u_j(y_k-x)$$

与 $u_j(y_k-x)\leqslant|y_k-x|=|y-x|(1\leqslant j\leqslant m,k=1,2)$ 推出 $u_i(y_k-x)=|y-x|(1\leqslant i\leqslant n,k=1,2)$. 由(IP)，必定 $y_1=y_2$. □

现在已可得出 Chebyshev 逼近定理的以下推广：

**5.11.8 定理** 设 $X=C(T)$，$T$ 为紧 $T_2$ 空间，$D=\text{span}\{e_1,\cdots,e_n\}\subset X$ 是 $n$ 维子空间，$x\in D^c$，$y\in D$，$z=y-x$. (i) $y\in Px\Leftrightarrow$ 存

在 $m(\leqslant n+1)$ 个不同点 $t_j\in T$ 与 $0\neq\alpha_j\in\mathrm{R}$，使 $\sum_j\alpha_j e_i(t_j)=0$ $(1\leqslant i\leqslant n)$，$|z(t_j)|=|z|_0$，$\alpha_j z(t_j)>0(1\leqslant j\leqslant m)$；若 $D$ 有(IP)，则 $y=Px$ 且 $m=n+1$. (ii)若 $T=J$，$e_i(t)=t^{i-1}(1\leqslant i\leqslant n)$，则 $y=Px$ $\Leftrightarrow$ 存在分点 $0\leqslant t_1<t_2<\cdots<t_{n+1}\leqslant 1$，使 $|z(t_j)|=|z|_0$，$z(t_j)=(-1)^{j-1}z(t_1)(1\leqslant j\leqslant n+1)$.

**证** (i)若 $y\in Px$，$\lambda_j,u_j,u$ 如 5.11.5，则 $u_j=\varepsilon_j\delta_{t_j}$，$\varepsilon_j=\pm 1$，$t_j$ 互不相同，$\delta_t$ 是集中于 $t$ 的 Dirac 测度. 令 $\alpha_j=\varepsilon_j\lambda_j$，则 $\alpha_j z(t_j)=\lambda_j|z|_0>0$. 逆命题不必证.

(ii)若 $y=Px$，则由(i)有 $0\leqslant t_1<t_2<\cdots<t_{n+1}=1$，$\alpha_j\neq 0$，使 $|z(t_j)|=|z|_0$，$\alpha_j z(t_j)>0$，

$$\sum_{j=1}^{n}\alpha_j t_j^{i-1}=-\alpha_{n+1}t_{n+1}^{i-1},\ i=1,2,\cdots,n. \tag{9}$$

用初等的论证从(9)得出 $\mathrm{sgn}\alpha_j=(-1)^{n+1-j}\mathrm{sgn}\alpha_{n+1}$，于是

$$z(t_j)/z(t_1)=\mathrm{sgn}\alpha_j/\mathrm{sgn}\alpha_1=(-1)^{j-1}(1\leqslant j\leqslant n+1). \quad \square$$

# 第六章　变分不等式

近年来已成为重要数学分支的变分不等式理论，有广泛的应用前景与相对独立的方法体系，非专著难以论其详. 本章只是立足于前几章所确立的理论框架之内，从单调算子与凸最优化理论演绎出变分不等式的若干基本结果，借以阐明抽象方法在这一领域所能起的作用.

本章总设 $X$ 是自反 Banach 空间（较精确的结果要求 $X$ 是 Hilbert 空间），$D\subset X$ 为非空闭凸集.

## §1　基本存在定理

变分不等式(VIE)的一般形式为

$$x \in D: \langle Fx, y - x\rangle \geqslant f(x) - f(y) (\forall\ y \in D), \qquad (1)$$

其中 $F: D\to X^*$，$f: D\to \mathrm{R}$. 当 $x\in D$ 满足(1)时，称 $x$ 为 VIE(1)的解. 若认定 $f|D^c=\infty$，则(1)中的"$y\in D$"可代以"$y\in X$"，于是 $x$ 是(1)的解等价于（参看 4.2.1）

$$x \in D: 0 \in Fx + \partial f(x). \qquad (2)$$

若取 $f=\delta_D$，则(1)与(2)分别成为

$$x \in D: \langle Fx, y - x\rangle \geqslant 0 \quad (\forall\ y \in D) \qquad (3)$$

与

$$x \in D: 0 \in Fx + \partial\delta_D(x). \qquad (4)$$

(3)是应用上最重要的形式. 若 $F$ 是单调位势算子，即有凸泛函 $\varphi$: $D\to\mathrm{R}$，使 $F=D\varphi$，则 $x$ 是(3)的解$\Leftrightarrow\varphi(x)=\min\limits_{D}\varphi(y)$（4.2.6）. 这显示出变分不等式与凸最优化问题有自然的联系.

$D$ 为闭凸锥与闭子空间这两种特殊情况是常见且重要的. 显然 $x\in D$ 是(3)的解$\Leftrightarrow\langle Fx, D-x\rangle\geqslant 0\Leftrightarrow Fx\in(D-x)^*$. 容易验证，当 $D$ 为闭凸锥[闭子空间]时 $(D-x)^*=D^*\cap\{x\}^\perp[(D-x)^*$

$=D^{\perp}$],于是有

**6.1.1 命题** 若 $D$ 为闭凸锥,则 $x$ 是 VIE(3)的解当且仅当

$$x \in D; Fx \in D^*, \langle Fx, x\rangle = 0. \tag{5}$$

若 $D$ 是闭子空间,则 $x\in D$ 是(3)的解$\Leftrightarrow Fx\in D^{\perp}$.

**6.1.2 命题** 若 $F$ 单调、半连续,$f$ 凸,则 $x\in D$ 是 VIE(1)的解当且仅当

$$\langle Fy, y-x\rangle \geqslant f(x)-f(y) \quad (\forall\ y\in D). \tag{6}$$

**证** 若 $x$ 是(1)的解,则由 $F$ 单调与(1)有

$$\langle Fy, y-x\rangle \geqslant \langle Fx, y-x\rangle \geqslant f(x)-f(y) \quad (\forall\ y\in D).$$

反之,若 $x\in D$ 满足(6),则在(6)中以 $x+t(y-x)$($t\in(0,1)$,$y\in D$)代 $y$ 并用 $f$ 的凸性得

$$\begin{aligned}\langle F(x+t(y-x)), y-x\rangle &\geqslant t^{-1}[f(x)-f(x+t(y-x))]\\ &\geqslant f(x)-f(y);\end{aligned}$$

令 $t\downarrow 0$ 得 $\langle Fx, y-x\rangle\geqslant f(x)-f(y)$,即(1)满足. □

因不等式(6)的解集可表为

$$\bigcap_{y\in D}\{x\in D; f(x)+\langle Fy, x\rangle\leqslant f(y)+\langle Fy, y\rangle\},$$

而 $f$ 凸[lsc]$\Rightarrow f+Fy$ 凸[lsc],故由 6.1.2 得出

**6.1.3 推论** 设 6.1.2 之条件满足,则 VIE(1)之解集是凸集,且当 $f$ lsc 时解集是闭凸集.

**6.1.4 命题** 若 $F$ 严格单调,则 VIE(1)至多有一解.

**证** 若 $x,z$ 是(1)的解,则必

$$\langle Fx, z-x\rangle \geqslant f(x)-f(z);$$

$$\langle Fz, x-z\rangle \geqslant f(z)-f(x),$$

两式相加得 $\langle Fx-Fz, z-x\rangle\geqslant 0$,这得 $x=z$. □

本节的主要课题是给出(1)与(3)的可解性条件.首先注意到,当 $f$ 凸且 lsc 时 $\partial f$ 极大单调(4.2.7),于是直接应用 3.3.1 于(2)得出:

**6.1.5 基本存在定理** 设 $F$ 伪单调、次连续、有界且满足如下强制性条件:

(C)$\exists\ \hat{x}\in D,\hat{u}\in\partial f(\hat{x})$，当 $x\in D,|x|\to\infty$ 时

$$\langle Fx+\hat{u},x-\hat{x}\rangle>0;$$

$f$ 凸且 lsc，则 VIE(1)有解.

**注** 条件"$F$ 伪单调、次连续"可代以更强的条件"$F:X\to X^*$ 单调半连续"(3.1.9)；当 $F=D\varphi,\varphi:X\to\mathrm{R}$ 凸时 $F$ 必单调半连续. 条件(C)可代以较强的条件"$\exists\ \hat{x}\in D:\partial f(\hat{x})\neq\varnothing$ 且 $F(\hat{x}+\cdot)$强制"；而后者又可代以更强的条件"$\exists\ \hat{x}\in D:\partial f(\hat{x})\neq\varnothing$，$F$ 强制且 $\overline{\lim\limits_{x\in D,|x|\to\infty}}|Fx|/|x|<\infty$"(参考 3.3.2，3.3.3).

因 $\delta_D$ 凸且 lsc，$0\in\partial\delta_D(x)(\forall\ x\in D)$，于是有

**6.1.6 推论** 设 $F$ 伪单调、次连续、有界且$\exists\ \hat{x}\in D:F(\hat{x}+\cdot)$强制，则 VIE(3)有解.

不依赖 3.3.1，可直接证明以下结果：

**6.1.7 定理** 设 $F$ 单调、次连续；$\exists\ \hat{x}\in D:F(\hat{x}+\cdot)$强制，则 VIE(3)有解.

**证** 1°设 $X=\mathrm{R}^n$，$D$ 有界，此时 $F:D\to\mathrm{R}^n$ 连续. 设 $P:X\to D$ 依 §5.11，对 $P(I-F)$用 Brouwer 不动点定理得 $x=P(x-Fx)\in D$. 由 5.11.1(iii)有

$$\begin{aligned}(Fx,D-x)&=(D-P(x-Fx),P(x-Fx)-(x-Fx))\\&\geqslant 0,\end{aligned}$$

可见 $x$ 是(3)的解.

2°设 $\dim X=\infty$，$D$ 有界. $\forall\ y\in D$，令 $S(y)=\{x\in D:\langle Fy,y-x\rangle\geqslant 0\}$，则 $S(y)$有界、弱闭，从而弱紧，$\bigcap\limits_{y\in D}S(y)$是(3)的解集(6.1.2). 任给有限集$\{y_i\}\subset D$，令 $Y=\mathrm{span}\{y_i\}$. 将第 1°段所证用到

$$F:D\cap Y\subset Y\to Y^*,x\mapsto Fx|Y,$$

知有 $x\in D\cap Y$，$\forall\ y\in D\cap Y:\langle Fy,y-x\rangle\geqslant 0$，于是 $x\in\bigcap S(y_i)$. 因$\{S(y):y\in D\}$是有限相交的弱紧集族，故$\bigcap\limits_{y\in D}S(y)\neq\varnothing$.

3°设 $D$ 无界，令 $D_n=D\cap\overline{B}_n(0)$. 由已证结论，当 $n$ 充分大时

$\exists\ x_n\in D_n:\langle Fx_n,D_n-x_n\rangle\geqslant 0$；可设 $\hat{x}\in D_n$，于是 $\langle Fx_n,x_n-\hat{x}\rangle\leqslant 0$，这表明 $\{x_n\}$ 有界. 取定 $n$，使 $n>|x_n|$. $\forall\ y\in D$，当 $t\downarrow 0$ 时 $y_t\triangleq x_n+t(y-x_n)\in D_n$，于是

$$\langle Fx_n,y-x_n\rangle=t^{-1}\langle Fx_n,y_t-x_n\rangle\geqslant 0,$$

这表明 $x_n$ 是(3)的解. □

注意 6.1.7 不要求 $F$ 有界；若 $D$ 有界，则可去掉强制性条件；若 $X=\mathrm{R}^n$，$D$ 有界，则只需 $F$ 连续就够了.

**6.1.8 推论** 设 $X$ 为 Hilbert 空间，$G:D\to D$，$\mathrm{Lip}G\leqslant 1$，$F=I-G$ 强制，则 $\mathrm{Fix}G$ 是非空有界闭凸集.

**证** $F$ 单调(3.7.3)、连续，且易见 $\forall\ \hat{x}\in D:F(\hat{x}+\cdot)$ 强制，于是由 6.1.7 知(3)的解集 $S$ 非空. 由 6.1.3 知 $S$ 闭凸；由 $F$ 强制易推出 $S$ 有界. 余下只需指明 $\mathrm{Fix}G=S$. 显然 $\mathrm{Fix}G=F^{-1}(0)\subset S$. 其次，$\forall\ x\in S$，由 $0\leqslant\langle Fx,Gx-x\rangle=-|Fx|^2$ 推出 $Fx=0$，从而 $x\in\mathrm{Fix}G$. □

## §2 二次变分不等式

本节考虑二次变分不等式

$$x\in D:\langle Ax-b,y-x\rangle\geqslant f(x)-f(y)(\forall\ y\in D)\qquad(1)$$

及其特款(取 $f=0$)

$$x\in D:\langle Ax-b,D-x\rangle\geqslant 0,\qquad(2)$$

其中 $A\in L(X,X^*)$，$b\in X^*$，$f:X\to\overline{\mathrm{R}}$，$D=D_f$. 令 $a(x,y)=\langle Ax,y\rangle$，则 $a:X\times X\to\mathrm{R}$ 是连续双线性函数. 通常将(1)的左端写成 $a(x,y-x)-\langle b,y-x\rangle$，它是 $x$ 的"二次函数".

二次变分不等式较为简单且应用广泛，因此特别受到注意，有关它的理论也比较完善. 首先，基于上节的一般理论容易建立：

**6.2.1 定理** 若 $A$ 强单调，$f$ 凸、lsc 且在某点 $\hat{x}\in D$ 次可微，则对每个 $b\in X^*$，VIE(1)有唯一解 $x=x(b)$，且 $b\mapsto x(b)$ 为 Lip.

**证** 解 $x(b)$ 的存在唯一性直接由 6.1.4，6.1.5 得出. 设 $x_i=$

$x(b_i), b_i \in X^* (i=1,2)$，则

$$
\begin{aligned}
0 &= [f(x_1) - f(x_2)] + [f(x_2) - f(x_1)] \\
&\leqslant \langle Ax_1 - b_1, x_2 - x_1 \rangle + \langle Ax_2 - b_2, x_1 - x_2 \rangle \\
&= \langle A(x_1 - x_2) - b_1 + b_2, x_2 - x_1 \rangle \\
&\leqslant -\lambda |x_1 - x_2|^2 + |b_1 - b_2| |x_1 - x_2|,
\end{aligned}
$$

其中 $\lambda$ 是正常数（注意 $A$ 强单调）. 由此得出 $|x_1 - x_2| \leqslant \lambda^{-1}|b_1 - b_2|$，这表明 $\mathrm{Lip}x(\cdot) \leqslant \lambda^{-1}$. □

对于 VIE(2)，结论可加强些.

**6.2.2 定理** 若 $A$ 单调，则 VIE(2)的解集是凸闭集，且当 $D$ 有界时(2)必有解. 若 $A$ 强单调，则(2)有唯一解 $x(b)$，当 $D$ 为闭凸锥[闭子空间]时 $b \mapsto x(b)$ 为正齐次[线性]连续函数.

**证** 定理的前半部分由 6.1.3，6.1.7 得出. 下面设 $A$ 强单调，于是(2)有唯一解 $x(b)$ 且 $b \mapsto x(b)$ 连续. 若 $D$ 为闭凸锥，$t \geqslant 0$，则由 6.1.1 有

$$
tAx(b) - tb = t[Ax(b) - b] \in D^*;
$$

$$
\langle tAx(b) - tb, tx(b) \rangle = t^2 \langle Ax(b) - b, x(b) \rangle = 0,
$$

由此得出 $x(tb) = tx(b)$，即 $x(\cdot)$是正齐次的. 类似地可证，当 $D$ 为闭子空间时 $x(\cdot)$是线性的. □

**6.2.3 例** 设 $X = L^2(\Omega)$，$(\Omega, \mu)$是一测度空间. 取定 $u \in X$，$D \triangleq \{x \in X : x \geqslant u\} \subset X$ 为闭凸集. 任给 $b \in X$，由 6.2.2，VIE

$$
x \in D : (x - b, D - x) \geqslant 0 \tag{3}
$$

有唯一解 $x(b)$. 令 $u \vee b = \max\{u, b\}$，直接验知

$$
\begin{aligned}
&(u \vee b - b, y - u \vee b) \\
&= \int_{u \geqslant b} (u - b)(y - u) \mathrm{d}\mu \geqslant 0 (\forall\, y \in D),
\end{aligned}
$$

故必 $x(b) = u \vee b$.

若仅假定 $A$ 单调，则 VIE(2)解的存在与唯一性都成问题. 在这种情况下，考虑如下逼近方法：取强单调的 $B \in L(X, X^*)$与 $v \in X^*$，$\forall\, \varepsilon > 0$，因 $A + \varepsilon B$ 强单调，由 6.2.2，VIE

$$x \in D: \langle Ax - b + \varepsilon(Bx - v), D - x\rangle \geqslant 0 \qquad (4)_\varepsilon$$

有唯一解 $x_\varepsilon$. 现在的问题是：在什么条件下，当 $\varepsilon \downarrow 0$ 时 $x_\varepsilon$ 收敛于(2)的解？

**6.2.4 定理** 设 $A$ 单调；$B \in L(X, X^*)$ 强单调，$v \in X^*$，$x_\varepsilon$ 是 VIE$(4)_\varepsilon$ 的解. 则 VIE(2)有解$\Leftrightarrow\{x_\varepsilon: \varepsilon > 0\}$有界；若$\{x_\varepsilon: \varepsilon > 0\}$有界，则当 $\varepsilon \downarrow 0$ 时 $x_\varepsilon$ 收敛于(2)的一个解.

**证** 以 $S$ 记 VIE(2)的解集. 若 $x \in S$，则

$$\begin{aligned} 0 &\leqslant \varepsilon^{-1}\langle Ax_\varepsilon - Ax, x_\varepsilon - x\rangle \\ &\leqslant \varepsilon^{-1}\langle Ax_\varepsilon - b, x_\varepsilon - x\rangle \leqslant \langle Bx_\varepsilon - v, x - x_\varepsilon\rangle \\ &\leqslant -\langle Bx_\varepsilon, x_\varepsilon\rangle + |B||x_\varepsilon||x| + |v||x - x_\varepsilon|, \end{aligned}$$

这结合 $B$ 强单调得出 $x_\varepsilon(\varepsilon > 0)$有界.

其次设 $x_\varepsilon$ 有界. 取 $\varepsilon_n \downarrow 0$，使$x_n \triangleq x_{\varepsilon_n} \rightharpoonup x_0$，则

$$\begin{aligned} \langle Ay - b, y - x_0\rangle &= \lim_n \langle Ay - b, y - x_n\rangle \\ &\geqslant \overline{\lim_n}\langle Ax_n - b, y - x_n\rangle \\ &= \overline{\lim_n}\langle Ax_n - b + \varepsilon_n(Bx_n - v), y - x_n\rangle \\ &\geqslant 0 (\forall\, y \in D), \end{aligned}$$

可见 $x_0 \in S$(6.1.2). 因 $S$ 非空闭凸(6.2.2)，VIE

$$x \in S: \langle Bx - v, S - x\rangle \geqslant 0 \qquad (5)$$

有唯一解 $\bar{x}$. 注意

$$\varepsilon_n\langle Bx_n - v, \bar{x} - x_n\rangle \geqslant \langle Ax_n - b, x_n - \bar{x}\rangle \geqslant 0. \qquad (6)$$

因凸函数 $x \mapsto \langle Bx, x\rangle$为 wlsc(4.1.6)，故从(6)得$\langle Bx_0 - v, \bar{x} - x_0\rangle \geqslant 0$. 这与$\langle B\bar{x} - v, x_0 - \bar{x}\rangle \geqslant 0$ 一起推出$\langle Bx_0 - B\bar{x}, \bar{x} - x_0\rangle \geqslant 0$，因此 $\bar{x} = x_0$. 设$\langle Bx, x\rangle \geqslant \lambda|x|^2(\forall\, x \in X)$，$\lambda > 0$，则

$$\begin{aligned} \overline{\lim_n}\lambda|x_n - \bar{x}|^2 &\leqslant \overline{\lim_n}\langle Bx_n - B\bar{x}, x_n - \bar{x}\rangle \\ &\leqslant \lim_n\langle v - B\bar{x}, x_n - \bar{x}\rangle = 0, \end{aligned}$$

可见 $x_n \to \bar{x}$. 因 $\bar{x}$ 与$\{\varepsilon_n\}$的选取无关，故 $x_\varepsilon \to \bar{x} \in S(\varepsilon \downarrow 0)$.

**6.2.5 推论** 设 $X$ 是 Hilbert 空间，$x_\varepsilon$ 是 VIE

$$x \in D:(Ax - b + \varepsilon x, D - x) \geqslant 0$$

的解，$A$ 单调. 若 VIE(2)有解，则当 $\varepsilon \downarrow 0$ 时 $x_\varepsilon$ 收敛于(2)的一个解.

作为强单调性的替代，引进所谓"半强制性". 设 $X$ 是 Hilbert 空间，$N=N(A)$，$N_b=N(b)$，分别以 $P$，$Q$ 记 $X$ 到 $N$ 与 $N\cap N_b$ 上的正投影算子，$P'=I-P$，$Q'=I-Q$. 关于 $A$ 的"半强制性"界定为：

(H) $\exists\ \lambda>0, \forall\ x\in X:(Ax,x)\geqslant\lambda|P'x|^2$.

**6.2.6 定理** 设 $X$ 是 Hilbert 空间，$A=A^*\in L(X)$ 单调且满足条件(H)，$\dim N<\infty$，$Q'(D)$闭. (i)若 $x,x'$ 是 VIE(2)的解，则 $x-x'\in N$. (ii)若 $D$ 是闭凸锥，则(2)有解$\Leftrightarrow(b,D\cap N)\leqslant 0$.

**证** (i)首先指明 $N=\{x:(Ax,x)=0\}$. 显然 $Ax=0\Rightarrow(Ax,x)=0$. 若$(Ax,x)=0$，则$\forall\ t\in\mathrm{R}, y\in X$，有

$$0\leqslant(A(x+ty),x+ty)=2t(Ax,y)+t^2(Ay,y),$$

由此得出 $Ax=0$. 若 $x,x'$ 是(2)的解，则如 6.1.4 之证有$(Ax-Ax',x'-x)\geqslant 0$，从而$(Ax-x',x-x')=0$，因此 $x-x'\in N$.

(ii)若(2)有解 $x$，则 $Ax-b\in D^*$(6.1.1)，于是对任给 $y\in D\cap N$ 有

$$0\leqslant(Ax-b,y)=(x,Ay)-(b,y)=-(b,y),$$

这表明$(b,D\cap N)\leqslant 0$.

其次设$(b,D\cap N)\leqslant 0$，今证(2)有解，为此只需证 $\varphi(x)\triangleq 2^{-1}(Ax,x)-(b,x)$在 $D$ 上取得极小. 取 $x_n\in D$，使 $\varphi(x_n)\to\alpha\triangleq\inf\varphi(D)$；令 $y_n=Q'x_n$. 若能证 $y_n\rightharpoonup y$，则因 $Q'(D)$凸闭(从而弱闭)，必定 $y\in Q'(D)$. 设 $y=Q'x, x\in D$，则

$$\varphi(x)=\varphi(y+Qx)=\varphi(y)\leqslant\varliminf_n\varphi(y_n)=\lim_n\varphi(x_n)=\alpha,$$

因此 $\varphi(x)=\min\varphi(D)$. 余下仅需证$\{y_n\}$有界.

若$\{y_n\}$无界，不妨设$|y_n|\to\infty$. 令 $z_n=x_n/|y_n|$，则由条件(H)有

$$\lambda|P'z_n|^2 \leqslant (Az_n, z_n) = 2|y_n|^{-2}[\varphi(x_n) + (b, x_n)]$$
$$= 2|y_n|^{-2}[\varphi(x_n) + (b, y_n)] \to 0, \tag{7}$$

于是 $P'z_n \to 0 (n \to \infty)$. 令 $T = P - Q$, 则 $Q' = T + P'$, $TX \cap P'X \subset N \cap N^{\perp} = \{0\}$, 因此有正交分解 $Q'X = TX \oplus P'X$, 从而

$$|Tz_n|^2 = |Q'z_n|^2 - |P'z_n|^2 = 1 - |P'z_n|^2 \to 1 (n \to \infty).$$

因 $Tz_n \in N$ 而 $\dim N < \infty$, 故不妨设 $Tz_n \to z \in N$, $|z| = 1$; 从而 $Q'z_n \to z \in Q'(D)$. 设 $z = Q'y$, $y \in D$, 则 $(b, z) = (b, y)$. 因 $y - z = Qy \in N$, $z \in N$, 故 $y \in D \cap N$, 因此 $(b, y) \leqslant 0$. 若 $(b, y) = 0$, 则 $y \in N \cap N_b$, 从而 $z = Q'y = 0$, 与 $|z| = 1$ 矛盾. 若 $(b, y) < 0$, 则 $(b, z) < 0$, 于是由(7)有

$$\varlimsup_n \frac{\lambda}{2}|y_n| |P'z_n|^2 \leqslant \varlimsup_n [|y_n|^{-1}\varphi(x_n) + (b, Q'z_n)]$$
$$= (b, z) < 0,$$

亦得矛盾. 因此 $\{y_n\}$ 有界. □

## §3 椭圆边值问题

在§1中已看到, 某些 VIE 问题可转化为方程问题. 反过来, 亦可考虑将方程问题转化为 VIE 问题, 对于下面讨论的椭圆边值问题就是如此.

设 $\Omega \subset \mathbb{R}^n$ 是有光滑边界的有界区域. 令 $X = H^1(\Omega)$, $Y = H_0^1(\Omega)$, 在 $X$ 与 $Y$ 中分别采用范数:

$$|u|_X = (|u|_2^2 + |\nabla u|_2^2)^{1/2}, |v|_y = |\nabla v|_2 (u \in X, v \in Y),$$

其中 $|\cdot|_2$ 记 $L^2$ 范数. 熟知 $|\cdot|_Y$ 满足 Poincáre 不等式:

$$|v|_2^2 \leqslant \beta |v|_Y^2 (\forall\, v \in Y), 0 < \beta = \text{const}. \tag{1}$$

首先考虑 2 阶 BVP:

$$\begin{cases} Lu \triangleq -\sum\limits_{i,j} \dfrac{\partial}{\partial x_i}\left(a_{ij}\dfrac{\partial u}{\partial x_j}\right) + cu = f, & \text{在 } \Omega \text{ 内;} \\ u = \varphi, & \text{在 } \partial\Omega \text{ 上,} \end{cases} \tag{2}$$

其中 $a_{ij},C\in L^{\infty}(\Omega),f\in H^{-1}(\Omega),\varphi\in X$. 若

$$m \triangleq \inf\left\{\sum_{i,j}a_{ij}(x)\xi_i\zeta_j : x\in\Omega,(\xi_i)\in S^{n-1}\right\}>0, \tag{3}$$

则说 $L$ 是椭圆的. 以 $(\cdot,\cdot)$ 记 $L^2$ 内积,令

$$a(u,v)=\sum_{i,j}\left(a_{ij}\frac{\partial u}{\partial x_j},\frac{\partial u}{\partial x_i}\right)+(cu,v), \tag{4}$$

则 $a:X\times X\to\mathrm{R}$ 是一连续双线性泛函,于是有唯一 $A\in L(X,X^*)$,使 $a(u,v)=\langle Au,v\rangle(u,v\in X)$. 熟知 $u\in X$ 是 BVP(2)的广义解(弱解)意味着(参照 § 3.6):

$$u-\varphi\in Y,a(u,v)=\langle f,v\rangle(\forall\, v\in Y). \tag{5}$$

令 $w=u-\varphi,b=f-A\varphi$,则(5)相当于:

$$w\in Y,\langle Aw-b,Y\rangle=0. \tag{6}$$

由 6.1.1,(6)恰好意味着 $w$ 是如下 VIE 的解:

$$w\in Y:\langle Aw-b,Y-w\rangle\geqslant 0. \tag{7}$$

于是应用 6.2.2 于 VIE(7)可建立:

**6.3.1 定理** 设条件(3)满足,$\lambda\triangleq\inf\limits_{\Omega}c(x)>-m/\beta$,$m,\beta$ 分别依(3)与(1),则 BVP(2)有唯一广义解.

**证** $\forall\, v\in Y$,由 $Av$ 及 $m,\lambda$ 的定义有

$$\langle Av,v\rangle=\sum_{i,j}\left(a_{ij}\frac{\partial v}{\partial x_i},\frac{\partial v}{\partial x_j}\right)+(cv,v)$$

$$\geqslant m|\nabla v|_2^2+\lambda|v|_2^2.$$

若 $\lambda\geqslant 0$,则 $\langle Av,v\rangle\geqslant m|v|_Y^2$;若 $0>\lambda>-m/\beta$,则由(1)有

$$\langle Av,v\rangle\geqslant(m+\lambda\beta)|v|_Y^2,m+\lambda\beta>0.$$

总之,当 $\lambda>-m/\beta$ 时 $A|Y$ 强单调,于是由 6.2.2 知(7)有唯一解,从而(2)有唯一广义解. □

6.3.1 特别可用于 BVP

$$-\Delta u+\lambda u=f(\text{在}\ \Omega\ \text{内}),u|\partial\Omega=\varphi. \tag{8}$$

对于 $Lu=-\Delta u+\lambda u$,依(3)有 $m=1$,于是由 6.3.1 有

**6.3.2 推论** 设 $f\in H^{-1}(\Omega),\varphi\in X,\lambda>-1/\beta$,则 BVP(8)有

唯一广义解(参看 3.6.4).

现在将类似的方法用到 Neumann 问题

$$\begin{cases} -\Delta u+\lambda u=f, & \text{在 } \Omega \text{ 内;} \\ \dfrac{\partial u}{\partial n}=\varphi, & \text{在 } \partial\Omega \text{ 上,} \end{cases} \tag{9}$$

其中 $f,\varphi$ 仍如 6.3.2,而 $\partial/\partial n$ 记沿 $\partial\Omega$ 的外法向导数.令

$$a(u,v)=(\nabla u,\nabla v)+\lambda(u,v); \tag{10}$$

$$\langle g,v\rangle=\langle f,v\rangle+\int_{\partial\Omega}\varphi v\mathrm{d}\sigma, \tag{11}$$

其中 $d\sigma$ 记 $\partial\Omega$ 的($n-1$ 维)面积元.直接看出 $g\in X^*$;而由前面的讨论知 $a(u,v)=\langle Au,v\rangle(\forall\, u,v\in X)$,$A\in L(X,X^*)$.约定 $u\in X$ 是 BVP(9)的广义解当且仅当

$$a(u,v)=\langle g,v\rangle(\forall\, v\in X). \tag{12}$$

从一个使用光滑资料的自然推理看来,以上约定是合理的.(12)意味着 $u$ 满足 VIE

$$u\in X:\langle Au-g,X-u\rangle\geqslant 0. \tag{13}$$

于是类似于 6.3.1 有

**6.3.3 定理** 若 $\lambda>0$,则 BVP(9)有唯一广义解.

**证** $\forall\, v\in X$,有

$$\langle Av,v\rangle=|\nabla v|_2^2+\lambda|v|_2^2\geqslant\min\{1,\lambda\}|v|_X^2,$$

可见 $A$ 强单调,于是由 6.2.2 得所要结论. □

**注** 若 $\lambda=0$,则 BVP(9)可能无解;即使有解也未必唯一.

(7)与(13)的约束集都是子空间,(13)的约束甚至是平凡的,这都不是 VIE 的典型情况,因此以上讨论还算不上 VIE 概念的本质应用.下面考虑一个带不等式约束的 BVP,它导致具有非子空间约束的 VIE,因而需要 VIE 概念的本质使用.

考虑问题(9)的以下变种:

$$\begin{cases} -\Delta u+\lambda u=f, & \text{在 } \Omega \text{ 内;} \\ u\geqslant 0,\ -\dfrac{\partial u}{\partial n}\geqslant\varphi,\left(\dfrac{\partial u}{\partial n}-\varphi\right)u=0, & \text{在 } \partial\Omega \text{ 上,} \end{cases} \tag{14}$$

其中 $f,\varphi$ 仍如 6.3.2. 称 $u\in X$ 为(14)的广义解,若

$$\begin{cases} a(u,u) = \langle g,u\rangle, u|\partial\Omega \geqslant 0; \\ a(u,v) \geqslant \langle g,v\rangle, \text{当 } v \in X, v|\partial\Omega \geqslant 0, \end{cases} \tag{15}$$

其中 $a,g$ 分别依(10)与(11). 令 $D=\{u\in X: u|\partial\Omega\geqslant 0\}$,则 $D\subset X$ 为闭凸锥. 对照 6.1.1 看出,(15)恰好意味着 $u$ 满足以下 VIE:

$$u \in D: \langle Au - g, D - u\rangle \geqslant 0.$$

于是如同 6.3.3 一样有:

**6.3.4 定理** 若 $\lambda>0$,则 BVP(14)有唯一广义解.

# §4 障碍问题

设 $\Omega,X,Y$ 如上节,定义(参看 §3(4))

$$a(u,v) = \sum_{i,j}\left(a_{ij}\frac{\partial u}{\partial x_j},\frac{\partial v}{\partial x_i}\right), u,v \in X, \tag{1}$$

其中 $a_{ij}\in L^\infty(\Omega)(1\leqslant i,j\leqslant n)$. 设 $a(u,v)=\langle Au,v\rangle, A\in L(X,X^*)$. 给定 $f\in H^{-1}(\Omega),\varphi\in X$,考虑 VIE:

$$u \in D: \langle Au - f, D - u\rangle \geqslant 0, \tag{2}$$

其中
$$D = \{u \in Y: u \geqslant \varphi\}. \tag{3}$$

若 $a_{ij}=a_{ji}$,则 $A=A^*$,于是 $Au-f=E'(u)$,此处 $E(u)=2^{-1}a(u,u)-\langle f,u\rangle(u\in X)$,因此当 $A$ 单调时 $u$ 是(2)的解$\Leftrightarrow E(u)=\min E(D)$(4.1.10). 特别,若 $A=-\Delta$,即 $a(u,v)=(\nabla u,\nabla v)$,则 $u\in D$ 是(2)的解$\Leftrightarrow u$ 使 $E(u)=2^{-1}|\nabla u|_2^2-\langle f,u\rangle(u\in D)$达到最小.

VIE(2)有如下物理解释:张于"曲面"$z=\varphi(x)$上的弹性薄膜周边固定在 $\partial\Omega$ 上,在外力作用下获得位移 $u$,当 $u$ 满足(2)时位能 $E(u)$最小. 因薄膜为 $z=\varphi(x)$所阻而需限定 $u\geqslant\varphi$,故此类问题被称为"障碍问题",它是导致变分不等式概念的典型问题之一.

$D$ 的定义式(3)涉及 $X$ 中序$\leqslant$的准确含义.

**6.4.1 定义** 任给 $S\subset\overline{\Omega},u\in X$,若存在 $u_k\in C^{1-0}(\overline{\Omega})$,使得 $u_k(x)\geqslant 0(\forall\ x\in S)$,$|u_k-u|_X\to 0(k\to\infty)$,则说在 $S$ 上 $u\geqslant 0$,写作

$u|S\geqslant 0$；$u|\Omega\geqslant 0(\Leftrightarrow u|\overline{\Omega}\geqslant 0)$就写作 $u\geqslant 0$.

易见$K_S\triangleq\{u\in X: u|S\geqslant 0\}$是 $X$ 中的闭凸锥，它在 $X$ 中导入一个序$\leqslant_S$. 约定$\sup\limits_S u=\inf\{\rho: u\leqslant_S\rho\}$.

**6.4.2 命题** 设 $S\subset\overline{\Omega}$，$u\in X$. (i)若 $u|S\geqslant 0$，则在 $S$ 上 $u(x)\geqslant 0$，a.e.；当 $S=\Omega$ 时其逆亦真. (ii)若 $u\in Y$，则 $u(x)\geqslant 0$，a.e. $\Leftrightarrow$ $\exists\ u_k\in Y\cap C^{1-0}(\overline{\Omega})$：$u_k\geqslant 0$ 且 $|u_k-u|_Y\to 0(k\to\infty)$.

证明是直接的.

以下设 $\varphi|\partial\Omega\leqslant 0$，从而 $\varphi^+=\varphi\vee 0\in D$，于是 $D\subset X$ 是非空闭凸集. 直接由 6.2.2 得出：

**6.4.3 定理** 若 $m>0$，$m$ 依 §3(3)，则 VIE(2)有唯一解.

任给 $u\in X$，$x\in\Omega$，若有 $x$ 的邻域 $V$ 与 $\varphi\in C_c^\infty(V)$，使得$(u-\psi)|V\geqslant 0$，$\psi(x)>0$，则说 $u(x)>0$；约定 $u(x)>\varphi(x)\Leftrightarrow(u-\varphi)(x)>0$，$u(x)=\varphi(x)\Leftrightarrow u\geqslant\varphi$ 且 $u(x)\not>\varphi(x)$. 若 $u$ 是 VIE(2)的解，则称

$$I(u)\triangleq\{x\in\Omega: u(x)=\varphi(x)\}\tag{4}$$

为 $u$ 的"贴合集". 障碍的作用体现于贴合集.

**6.4.4 定理** 若 $u$ 的 VIE(2)的解，则存在 $\Omega$ 上的正测度 $\mu$，使得 $\operatorname{supp}\mu\subset I(u)$，且

$$\langle Au-f,v\rangle=\int_\Omega v\mathrm{d}\mu(\forall\ v\in Y).\tag{5}$$

**证** 若 $0\leqslant v\in Y$，则 $u+v\in D$，于是

$$\langle Au-f,v\rangle=\langle Au-f,v+u-u\rangle\geqslant 0,$$

可见 $Au-f$ 是 $Y$ 上的正线性泛函. 由 Riesz 表示定理[75;3.4.5](注意 $C_c(\Omega)\subset Y$!)，存在 $\Omega$ 上的正测度 $\mu$，使(5)成立. 余下只要证 $\operatorname{supp}\mu\subset I(u)$，为此只要证在开集 $\Omega\backslash I(u)$内 $Au-f$(作为分布)为零. 任给 $x_0\in\Omega\backslash I(u)$，取 $r>0$，$\psi\in C_c^\infty(B_r(x_0))$，使 $\psi(x_0)>0$ 且$(u-\varphi-\psi)|B_r(x_0)\geqslant 0$. $\forall\ v\in C_c^\infty(B_r(x_0))$，当 $\varepsilon>0$ 充分小时在 $B_r(x_0)$上 $u\pm\varepsilon v\geqslant\varphi+\psi/2\geqslant\varphi$，于是

$$0\leqslant\langle Au-f,u\pm\varepsilon v-u\rangle=\pm\varepsilon\langle Au-f,v\rangle,$$

这推出$\langle Au-f,v\rangle=0$. 因此在$\Omega\backslash I(u)$内 $Au-f=0$. □

6.4.4 之结论可表为更富启发性的以下形式：

$$\begin{cases} Lu = f + \mu, & \text{在 } \Omega \text{ 内}; \\ Lu = f, & \text{在 } \Omega\backslash I(u) \text{ 内}, \end{cases} \tag{6}$$

其中

$$Lu = -\sum_{i,j} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right). \tag{7}$$

若$I(u)=\varnothing$，则 $u$ 是 BVP

$$Lu = f(\text{在 } \Omega \text{ 内}), u|\partial\Omega = 0 \tag{8}$$

的广义解. 可以说，$I(u)$体现了 VIE(2)与 BVP(8)的差别.

若 $u\in X$，当 $0\leqslant v\in Y$ 时$\langle Au-f,v\rangle\geqslant 0$，则说 $u$ 是 $L$-$f$ 上解，$L$ 依(7).

**6.4.5 定理** 设 $m>0$，$m$ 依 §3(3)，$u$ 是 VIE(2)的解，$v$ 是 $L$-$f$ 上解，$v\geqslant\varphi$，$v|\partial\Omega\geqslant 0$，则 $u\leqslant v$.

**证** 令 $w=\min\{u,v\}$，只需证 $w=u$. 由 $v|\partial\Omega\geqslant 0$ 推出 $w|\partial\Omega=0$，即 $w\in Y$，从而 $w\in D$，于是

$$\langle Au - f, w - u\rangle \geqslant 0. \tag{9}$$

由 $0\leqslant u-w\in Y$ 推出$\langle Av-f,u-w\rangle\geqslant 0$；结合(9)得$\langle Au-Av,w-u\rangle\geqslant 0$. 另一方面，

$$\begin{aligned} \langle Av - Au, w - u\rangle &= \sum_{i,j}\left(a_{ij}\frac{\partial(v-u)}{\partial x_j}, \frac{\partial(w-u)}{\partial x_i}\right) \\ &= \sum_{i,j}\int_{\Omega(v<u)} a_{ij}\frac{\partial(v-u)}{\partial x_j}\frac{\partial(v-u)}{\partial x_i}\mathrm{d}x \\ &= a(w-u, w-u) \geqslant m|w-u|_Y^2, \end{aligned}$$

因此有 $w=u$. □

**6.4.6 推论** 若 $u$ 是 VIE(2)的解而(2)中 $f=0$，$m>0$，$m$ 依 §3(3)；$\beta\triangleq\sup\limits_{\Omega}\varphi\geqslant 0$，则 $u\leqslant\beta$.

**证** 只需注意 $v\equiv\beta$ 是 $L$-0 上解. □

# 第七章　临界点理论

任给 $f \in C^1(X)$，称"Euler 方程"$f'(x)=0$ 的解为 $f$ 的临界点；临界点可能为极值点. 鉴于 Euler 方程在数学物理问题中有普遍意义，无论临界点是否为极值点，都有重要研究价值. 临界点理论的基本课题是：形成系统的方法，用以判明给定泛函的临界点是否存在，可能时估计临界点的个数并指明其分布状况. 这些问题的解决依赖于对"水平集"$f_a=\{x: f(x)\leqslant a\}$ 的考察，基本原理是：若 $a$ 的变化引起 $f_a$ 的拓扑结构发生变化，则意味着临界点出现. 鉴于问题的拓扑本质，一定拓扑工具的使用是不可避免的.

## §1　Banach 流形

临界点概念本身是局部的，但临界点的分布却是一个典型的大范围分析问题，因此只有在无限维流形的框架内才能得到完整的处理. 本节概述有关 Banach 流形的基本用语，它们在现代分析中是标准与通用的(参看[108，201，203]).

**7.1.1 定义**　设 $M$ 是一仿紧 Hausdorff 空间. 若存在一集 $\Phi=\{(U_\alpha,\varphi_\alpha): \alpha\in A\}$，使得：(i)$U_\alpha\subset M$ 开且 $M=\bigcup U_\alpha$；(ii)$\varphi_\alpha$ 是从 $U_\alpha$ 到某个 Banach 空间 $X_\alpha$ 的开子集 $\varphi_\alpha U_\alpha$ 上的同胚；(iii)$\varphi_\beta\varphi_\alpha^{-1}\in C^r(\alpha,\beta\in A)$，则称 $M$ 或$(M,\Phi)$为以 $\Phi$ 为图册的 $C^r$Banacah 流形，称

$$\Psi=\{(V,\psi): \Phi\cup\{(V,\psi)\}\text{ 满足条件(ii)(iii)}\}$$

为 $\Phi$ 生成的 $C^r$ 结构，称每个$(V,\psi)\in\Psi$ 为 $M$ 上的图，当 $x\in V$ 时说$(V,\psi)$是 $M$ 在 $x$ 处的图.

以下总假定 $r\geqslant1$. 当 $M$ 连通时，可指明 7.1.1 中的空间 $X_\alpha$ 互相同构. 为简单起见，假定 $X_\alpha\equiv X$，此时称 $M$ 为 $X$-流形；称 $\mathrm{R}^n$-流形为

$n$ 维流形. 显然 $X$ 本身是一 $C^\infty$ 流形,$(X,\mathrm{id})$即为其上的图.

以下设 $M,N$ 分别为 $C^r$ 的 $X$-流形与 $Y$-流形.

**7.1.2 定义** 设 $f:M\to N$. 若$\forall\ x\in M$,存在 $M$ 上的图$(U,\varphi)$与 $N$ 上的图$(V,\psi)$,使 $x\in U$,$fU\subset V$,$\psi f\varphi^{-1}\in C^r$,则称 $f$ 为 $C^r$ 映射,记作 $f\in C^r(M,N)$. 若 $f\in C^r(M,N)$且 $f^{-1}\in C^r(N,M)$,则称 $f$ 为 $C^r$ 同胚;约定 $\mathrm{Diff}^r(M)$记 $M$ 到自身的 $C^r$ 同胚之全体.

显然,$f:M\to Y$ 为 $C^r$ 映射$\Leftrightarrow\forall\ x\in M$,存在 $M$ 在 $x$ 处的图$(U,\varphi)$;$f\varphi^{-1}\in C^r$;$(U,\varphi)$是 $M$ 上的图$\Leftrightarrow U\subset M$ 与 $\varphi U\subset X$ 为开集且 $\varphi:U\to\varphi U$ 为 $C^r$ 同胚.

设$(U,\varphi)$与$(V,\psi)$是 $M$ 在 $x$ 处的图,$\xi,\eta\in X$. 约定$(x,\varphi,\xi)\sim(x,\psi,\eta)\Leftrightarrow\eta=(\psi\varphi^{-1})'(\varphi x)\xi$. 以$[x,\varphi,\xi]$记含$(x,\varphi,\xi)$的$\sim$等价类,其全体记为 $T_xM$,称为 $M$ 在 $x$ 的切空间,称 $TM\triangleq\bigcup\limits_{x\in M}T_xM$ 为 $M$ 的切丛,称映射 $\pi:TM\to M$,$[x,\varphi,\xi]\mapsto x$ 为丛投影. 在 $T_xM$ 中采用由双射 $\xi\mapsto[x,\varphi,\xi]$诱导的拓扑向量空间结构(它与 $\varphi$ 的选取无关),于是 $T_xM$ 是拓扑同构于 $X$ 的拓扑向量空间且可度量化为 Banach 空间. $TM$ 则可依自然的方式定义为一个 $C^{r-1}$类的 $X^2$-流形,其构造细节不必考虑.

**7.1.3 定义** 设 $f\in C^r(M,N)$,$x\in M$,$y=f(x)$. 称映射

$$df_x:T_xM\to T_yN,[x,\varphi,\xi]\mapsto[y,\psi,(\psi f\varphi^{-1})'(\varphi x)\xi]\qquad(1)$$

为 $f$ 在 $x$ 的导数,亦记作 $df(x)$. 由 $df|T_xM=df_x(\forall\ x\in M)$决定一映射 $df:TM\to TN$,称之为 $f$ 的切映射.

可验证(1)与 $\varphi,\psi$ 的选取无关,且 $df_x\in L(T_xM,T_yN)$,$df\in C^{r-1}(TM,TN)$. 当 $df_x$ 非满射时称 $x$ 为 $f$ 的临界点,称 $f(x)$为 $f$ 的临界值;若 $y\in N$ 不是 $f$ 的临界值,则称 $y$ 为 $f$ 的正则值. 若 $df_x$ 为满射且是核裂的,则说 $f$ 在 $x$ 为浸满(参看 5.3.3);$f$ 为浸满$\Leftrightarrow f$ 在每点 $x\in M$ 为浸满. 若$\forall\ x\in M$;$df_x$ 为单射且是值裂的,则称 $f$ 为浸入.

称 $T_x^* M \triangleq (T_x M)^*$ 为 $M$ 在 $x$ 的余切空间. 若 $f \in C^r(M) (\triangleq C^r(M, \mathrm{R}))$, 则 $\mathrm{d}f_x \in T_x^* M (x \in M)$; $x$ 是 $f$ 的临界点 $\Leftrightarrow \mathrm{d}f_x = 0$. 若 $g \in C^r(I, M)$, $I$ 是实区间, 则称 $g$ 为 $M$ 上的 $C^r$ 曲线, 称 $g'(t) \triangleq \mathrm{d}g_t(\mathrm{d}/\mathrm{d}t)(t \in I)$ 为 $g$ 的切向量, 此处 $\mathrm{d}/\mathrm{d}t$ 记 R 中的单位向量.

若 $f \in C^r(M, N)$, $g \in C^r(N, N')$, 则可验证"链规则":

$$\mathrm{d}(g \circ f)(x) = \mathrm{d}g(f(x)) \circ \mathrm{d}f(x), x \in M, \tag{2}$$

或简写作 $\mathrm{d}(g \circ f) = \mathrm{d}g \circ \mathrm{d}f$.

设 $M \subset N$, $i: M \to N$ 是包含映射. 若 $i$ 为 $C^r$ 浸入, 则称 $M$ 为 $N$ 的 $C^r$ 浸入子流形; 若 $i$ 是 $C^r$ 浸入且为拓扑嵌入(即 $M$ 是 $N$ 的拓扑子空间), 则称 $i$ 为 $C^r$ 嵌入且称 $M$ 为 $N$ 的 $C^r$ 子流形. 若 $M$ 是 $N$ 的 $C^r$ 子流形, 则 $\forall\, a \in M$, 有 $N$ 在 $a$ 处的图 $(V, \psi)$ 与分解 $Y = Y_1 \oplus Y_2$, 使 $\psi(M \cap V)$ 是 $Y_1$ 中的开集. 可证明若 $N$ 的子集 $M$ 有上述性质, 则必可定义 $M$ 为 $N$ 的 $C^r$ 子流形. 若 $i: M \to N$ 是 $C^r$ 浸入, $x \in M$, 则 $\mathrm{d}i_x: T_x M \to T_x N$ 是连续嵌入, 通常等同 $T_x M$ 与 $\mathrm{d}i_x(T_x M)$ 并认定 $T_x M \subset T_x N$. 若 $A, B$ 是 $M$ 的 $C^r$ 浸入子流形, $\forall\, x \in A \cap B$: $T_x M = T_x A + T_x B$, 则说 $A$ 与 $B$ 横截相交. $M$ 的开子集必为 $M$ 的 $C^r$ 子流形.

子流形经常由一定的映射产生, 以下是一标准的结果(通常称为"核与象定理").

**7.1.4 定理** 设 $f \in C^r(M, N)$. (i) 若 $y \in f(M)$, $f$ 在每点 $x \in f^{-1}y$ 为浸满, 则 $f^{-1}y$ 是 $M$ 的 $C^r$ 子流形, 且 $T_x(f^{-1}y) = N(\mathrm{d}f_x) (x \in f^{-1}y)$. (ii) 若 $f$ 是 $C^r$ 嵌入(即 $f$ 是 $C^r$ 浸入且为拓扑嵌入), 则 $f(M)$ 是 $N$ 的 $C^r$ 子流形, $f: M \to f(M)$ 是 $C^r$ 同胚, 且 $T_{f(x)} f(M) = \mathrm{d}f_x(T_x M) (x \in M)$.

证明是隐函数定理的标准应用, 从略.

7.1.4 特别用到 $f \in C^r(M)$ 的情况. 若 $y$ 是 $f$ 的正则值且 $A \triangleq f^{-1}(y) \neq \varnothing$, 则 $f$ 在 $A$ 上必为浸满, 因此 $A$ 是 $M$ 的 $C^r$ 子流形; $\forall\, x \in A$, $T_x A = N(\mathrm{d}f_x)$ 是 $T_x M$ 的极大子空间, 称这样的 $A$ 为 $M$ 中的超

曲面. 例如，设 $X$ 是 Hilbert 空间，$f(x)=|x|$，则 $f\in C^\infty(X)$，1 是 $f$ 的正则值，于是 $S\triangleq f^{-1}(1)$ 是 $X$ 中的超曲面. $\forall x\in S$，直接算出 $f'(x)=x$，于是 $T_xS=\{x\}^\perp$，这正是球面 $S$ 在 $x$ 的"切平面".

利用 7.1.4 可建立以下乘子结果(参照 5.2.7).

**7.1.5 定理** 设 $f\in C^r(X)$，$h\in C^r(X,Z)$，$M=h^{-1}(0)\neq\varnothing$，$h$ 在 $M$ 上为浸满. 则 $M$ 是 $X$ 的 $C^r$ 子流形，$f|M\in C^r$；$x\in M$ 是 $f|M$ 的临界点$\Leftrightarrow\exists\ \mu\in Z^*$：$f'(x)=\mu\circ h'(x)$.

**证** 由 7.1.4，$M$ 是 $X$ 的 $C^r$ 子流形，因此包含映射 $i:M\to X$ 是 $C^r$ 的，于是 $f|M=f\circ i\in C^r$. 设 $x\in M$，则 $x$ 是 $f|M$ 的临界点$\Leftrightarrow \mathrm{d}(f|M)_x=\mathrm{d}f_x\circ \mathrm{d}i_x=0\Leftrightarrow \mathrm{d}f_x|T_xM=0\Leftrightarrow T_xM=N(h'(x))\subset N(f'(x))\Leftrightarrow f'(x)\in N(h'(x))^\perp=R(h'(x)^*)\Leftrightarrow\exists\ \mu\in Z^*$：$f'(x)=\mu\circ h'(x)$. □

**7.1.6 定义** 若 $\xi\in C^r(M,TM)$，$\pi\circ\xi=\mathrm{id}$，则称 $\xi$ 为 $M$ 上的 $C^r$ 向量场. 若 $g:I\to M$ 是 $M$ 上的 $C^1$ 曲线，$0\in I$，$g(0)=x$，$g'(t)=\xi(g(t))$，则称 $g$ 为 $\xi$ 过点 $x$ 的积分曲线.

在本章中，与所考虑的泛函相关的一定向量场起重要作用. 求给定向量场 $\xi$ 的积分曲线，相当于解流形上的自治微分方程 $x'=\xi(x)$(通常称为"场微分方程")，在局部范围内，问题可转化为解一个 Banach 空间中的自治微分方程，而这有成熟的标准结论可用.

**7.1.7 定义** 设 $|\cdot|:TM\to\mathrm{R}_+$ 是一连续函数；$\forall\ x\in M$，$|\cdot|$ 在 $T_xM$ 上的限制 $|\cdot|_x$ 是一与 $T_xM$ 中的拓扑相容的范数；$\forall\ k>1$，$x\in M$，存在 $M$ 在 $x$ 的邻域 $U$ 与同胚 $TU\to U\times T_xM$，使得 $\forall\ y\in U$：$T_yM\to\{y\}\times T_xM$ 是同构，且当 $\eta\in T_yM$ 对应 $(y,\xi)\in\{y\}\times T_xM$ 时 $K^{-1}|\eta|_y\leqslant|\xi|_x\leqslant k|\eta|_y$，则称 $|\cdot|$ 为 $M$ 上的 Finsler 结构；称定义了 Finsler 结构的 Banach 流形为 Finsler 流形.

可以证明，7.1.1 意义下的 Banach 流形上恒可定义 Finsler 结构. 若 $M$ 是 $X$(或任何 Finsler 流形)的 $C^r$ 子流形，则每个 $T_xM(x\in M)$ 上有自然定义的范数 $|\cdot|_x$，且由此得出 $M$ 上的 Finsler 结构. 本

章并不用到 Finsler 结构的具体构成.

若 $M$ 是一连通的 Finsler 流形,$x,y\in M$,取 $M$ 上的 $C^1$ 曲线 $g(t)(a\leqslant t\leqslant b)$,使 $g(a)=x,g(b)=y$,称$L_g\triangleq\int_a^b|g'(t)|\mathrm{d}t$ 为 $g$ 的长度.定义

$$d(x,y)=\inf\{L_g:g\text{ 是 }M\text{ 上从 }x\text{ 至 }y\text{ 的 }C^1\text{ 曲线}\},\qquad(3)$$

则可证明 $d$ 是 $M$ 上的度量且与 $M$ 的拓扑相容.当 $d$ 为完备度量时称 $M$ 为完备的 Finsler 流形.

在本章中,总假定 $M$ 是给定的 $C^r$ 连通 $X$-流形,$r$ 大到足以满足行文的需要;$M$ 上已给定 Finsler 结构$|\cdot|$,且$|\cdot|$已导出一完备度量 $d$.

## §2 伪梯度场

给定 $f\in C^1(M)$,以 $K$ 记 $f$ 的临界点之全体,令 $\tilde{M}=M\backslash K$(本章将一直保持这些记号),显然 $\tilde{M}$ 是 $M$ 的开子流形.临界点理论的基本问题是:$f$ 是否有临界点?倘有临界点其个数与分布如何?这些问题的解答有赖于对"水平集"$f_a\triangleq\{x:f(x)\leqslant a\}$的考察.粗略地说,若 $a$ 变化时 $f_a$ 越过临界点,则 $f_a$ 的拓扑构造必随之改变.因此考察当 $a$ 递减时 $f_a$ 之嬗变规律可望获得临界点的信息.若 $M=X$ 而 $X$ 是 Hilbert 空间,则考察 $f_a$ 沿负梯度方向之形变是最自然的.对于一般的 Banach 流形 $M$,"梯度"$df$ 不再定义出 $M$ 上的向量场.幸而,有 Palais 发现的所谓"伪梯度"可以替代.

**7.2.1 定义** 若 $\tilde{M}$ 上的 $C^{1-0}$向量场 $\xi$ 满足

$$4^{-1}|\xi(x)|^2\leqslant|\mathrm{d}f(x)|^2\leqslant\langle\mathrm{d}f(x),\xi(x)\rangle(x\in\tilde{M}),\qquad(1)$$

则称 $\xi$ 为 $f$ 的伪梯度场.

不等式(1)蕴涵

$$0<2^{-1}|\xi(x)|\leqslant|\mathrm{d}f(x)|\leqslant\xi(x)(x\in\tilde{M});\qquad(2)$$

$$\langle \mathrm{d}f(x),\xi(x)\rangle/|\mathrm{d}f(x)||\xi(x)| \geqslant 1/2(x \in \widetilde{M}). \tag{3}$$

直观上(2)(3)可解释为：就模与"方向"两者而言，$\xi(x)$与$\mathrm{d}f(x)$的偏差都保持在一适度的范围内. 注意在 Hilbert 空间中，$\xi=f'$ 使(1)自动满足.

以下结果使得采用伪梯度是可行的.

**7.2.2 定理**(Palais[133],1966)　$f$ 的伪梯度场存在.

**证**　就 $M=X$ 的情况给出证明(一般情况的证明并无实质不同，只是记号更复杂些).

1°局部构成. $\forall\ x\in\widetilde{X}$，取 $\zeta(x)\in X$，使 $|\zeta(x)|=1$，$\langle f'(x),\zeta(x)\rangle>(2/3)|f'(x)|$. 令 $\eta(x)=(3/2)|f'(x)|\zeta(x)$，则易验知

$$4^{-1}|\eta(x)|^2<|f'(x)|^2<\langle f'(x),\eta(x)\rangle.$$

由连续性，有 $x$ 的开邻域 $V_x\subset\widetilde{X}$，使得

$$4^{-1}|\eta(x)|^2<|f'(y)|^2<\langle f'(y),\eta(x)\rangle(\forall\ y\in V_x). \tag{4}$$

2°整体构成. 由 $\widetilde{X}$ 仿紧，有 $\widetilde{X}$ 的局部有限开覆盖$\{U_\alpha\}$，使得$\forall\ \alpha$，$\exists\ x_\alpha\in\widetilde{X}$：$U_\alpha\subset V_{x_\alpha}$. 令 $g_\alpha(x)=d(x,U_\alpha^c)$，则 $\mathrm{Lip}g_\alpha\leqslant 1$，$g_\alpha|U_\alpha^c=0$. 令 $g=\sum g_\alpha$，则因$\sum g_\alpha$局部地为有限和，必定 $g\in C^{1-0}(\widetilde{X})$；其次显然 $g>0$. 令 $h_\alpha=g_\alpha/g$，则 $h_\alpha\in C^{1-0}(\widetilde{X},J)$，$\sum h_\alpha=1$. 定义

$$\xi(x)=\sum_\alpha h_\alpha(x)\eta(x_\alpha),x\in\widetilde{X}. \tag{5}$$

任给 $x\in\widetilde{X}$，若 $h_\alpha(x)\neq 0$，则 $x\in U_\alpha\subset V_{x_\alpha}$. 这一方面说明(5)之右端局部地为有限和，从而 $\xi\in C^{1-0}(\widetilde{X},X)$；另一方面，结合(4)得出：

$$\begin{aligned}4^{-1}|\xi(x)|^2 &\leqslant [2^{-1}\sum h_\alpha(x)|\eta(x_\alpha)|]^2\\ &\leqslant [\sum h_\alpha(x)|f'(x)|]^2\\ &=|f'(x)|^2=\sum h_\alpha(x)|f'(x)|^2\\ &\leqslant \sum h_\alpha(x)\langle f'(x),\eta(x_\alpha)\rangle\\ &=\langle f'(x),\xi(x)\rangle,\end{aligned}$$

这表明 $\xi$ 是 $f$ 的伪梯度场. □

$f$ 的伪梯度场显然不必是唯一的. 在本章中,所需要的是伪梯度场存在这一事实,而其具体构造并不重要.

**7.2.3 定理** 设 $\xi$ 是 $M$ 上的 $C^{1-0}$ 向量场,则存在开集 $W\subset \mathrm{R}\times M$ 与 $\sigma\in C(W,M)$, $W=\bigcup\limits_{x\in M}I_x\times\{x\}$, $I_x=(\alpha_x,\beta_x)$, $-\infty\leqslant\alpha_x<0<\beta_x\leqslant\infty$,使得 $\forall\ x\in M$: $\sigma(\cdot,x)$ 是 $\xi$ 过 $x$ 的极大积分曲线,且 $I_{\sigma(t,x)}=I_x-t$, $\sigma(s+t,x)=\sigma(s,\sigma(t,x))$ $(t,s+t\in I_x)$. 若 $\sup\limits_x|\xi(x)|<\infty$,则 $W=\mathrm{R}\times M$.

**证** 任给 $x_0\in M$,取 $M$ 在 $x_0$ 处的图 $(U,\varphi)$. 认定 $T(\varphi U)=\varphi U\times X$,则 $\xi_\varphi\triangleq d\varphi\circ(\xi|U)\circ\varphi^{-1}\in C^{1-0}(\varphi U,X)$. 将标准的微分方程结论用于 IVP

$$y'=\xi_\varphi(y),y(0)=\bar{y}\in\varphi U,\tag{6}$$

得出 $\theta\in C(I_\delta\times V_1,\varphi U)$, $I_\delta=(-\delta,\delta)$, $\delta>0$, $V_1$ 是 $\varphi(x_0)$ 的一邻域, $\forall\ \bar{y}\in V_1$, $y=\theta(\cdot,\bar{y})$ 是(6)在 $I_\delta$ 内的唯一解. 令 $V=\varphi^{-1}V_1$, $\sigma=(t,x)=\varphi^{-1}\theta(t,\varphi(x))$,则 $\sigma\in C(I_\delta\times V,U)$, $\sigma(0,x)=x$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma(t,x)=\mathrm{d}\varphi^{-1}\xi_\varphi(\theta(t,\varphi(x)))=\xi(\sigma(t,x)),$$

可见 $\sigma(\cdot,x)$ 是 $\xi$ 过 $x$ 的积分曲线,且必是 $\xi$ 过 $x$ 定义于 $I_\delta$ 上的唯一积分曲线.

任给 $x\in M$,由上段所证知有 $\xi$ 过 $x$ 的积分曲线 $\sigma_x$,设其极大定义区间为 $I_x=(\alpha_x,\beta_x)$, $-\infty\leqslant\alpha_x<0<\beta_x\leqslant\infty$. 令 $W=\bigcup\limits_{x\in M}I_x\times\{x\}$, $\sigma(t,x)=\sigma_x(t)$ $((t,x)\in W)$,则得映射 $\sigma:W\to M$, $\sigma(\cdot,x)$ 是 $\xi$ 过 $x$ 的极大积分曲线. 用上段的唯一性结论不难得出 $\sigma(s+t,x)=\sigma(s,\sigma(t,x))$ $(t,s+t\in I_x)$, $I_{\sigma(t,x)}=I_x-t$. 稍费事的是验证 $W$ 开且 $\sigma\in C(W,M)$. 令

$$W_0=\{(t,x):\text{有}(t,x)\text{ 的邻域 }W_1\subset W\text{ 使 }\sigma|W_1\text{ 连续}\},$$

只需证 $W_0=W$. 由上段所证知 $\{0\}\times M\subset W_0$. 若 $W_0\neq W$,则有 $(t_0,x_0)\in W\backslash W_0$,不妨设 $t_0=\inf\{t>0:(t,x_0)\overline{\in}W_0\}$,必定 $t_0>0$. 取 $x_1\triangleq$

$\sigma(t_0,x_0)$的邻域 $V$ 与 $\delta>0$，使 $I_\delta\times V\subset W_0$. 设 $t_0-\delta/3<t_1<t_0$，$x_2\triangleq\sigma(t_1,x_0)\in V$. 因 $(t_1,x_0)\in W_0$，故有 $x_0$ 的邻域 $V_0$，使 $t_1\times V_0\subset W_0$ 且 $\sigma(t_1,V_0)\subset V$，从而 $(t_0,x_0)\in(t_1+I_\delta)\times V_0\subset W_0$，得出矛盾.

若 $\mu\triangleq\sup\limits_x|\xi(x)|<\infty$，则当 $\alpha_x<\tau<t<\beta_x$ 时

$$\mathrm{d}(\sigma(t,x),\sigma(\tau,x))\leqslant\int_\tau^t\left|\frac{\mathrm{d}}{\mathrm{d}s}\sigma(s,x)\right|\mathrm{d}s\leqslant\mu(t-\tau).\qquad(7)$$

由此推出 $\beta_x=\infty$（否则(7)推出 $\sigma(\beta_x-0,x)$存在，与 $I_x$ 的极大性矛盾)；同理 $\alpha_x=-\infty$，因此 $W=\mathrm{R}\times M$. □

通常称 7.2.3 中的 $\sigma$ 为 $\xi$ 决定的流，称 $\sigma(\cdot,x)$为 $\xi$ 过 $x$ 的流线. 若 $W=\mathrm{R}\times M$，则$\sigma_t\triangleq\sigma(t,\cdot)\in\mathrm{Homeo}(M)$，$\sigma_{t+s}=\sigma_t\sigma_s(t,s\in\mathrm{R})$.

若 $\xi$ 是 $f$ 的伪梯度场，$\sigma$ 是一$\xi$ 的流，则

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}f(\sigma(t,x))&=-\langle\mathrm{d}f(\sigma(t,x)),\xi(\sigma(t,x))\rangle\\&\leqslant-|\mathrm{d}f(\sigma(t,x))|^2<0,\end{aligned}\qquad(8)$$

可见 $f$ 沿流线 $\sigma(\cdot,x)(x\in\widetilde{M})$严格下降，这正是本章要利用的关键事实. 不过，要使 $f$ 下降不致太慢，须对 $f$ 附加某种"紧性条件".

**7.2.4 定义** 若从 $f(x_n)$有界且$|\mathrm{d}f(x_n)|\to0$ 恒推出$\{x_n\}$含收敛子列，则说 $f$ 满足 Palais-Smale 条件(以下简称为(PS)).

**7.2.5 命题** 设 $f$ 满足(PS)，$a\leqslant b$，$K_{ab}=K\cap f^{-1}[a,b]\subset U$，$U\subset M$ 开. 则 $K_{ab}$紧；$\exists\ r>0$，使

$$\inf\{|\mathrm{d}f(x)|:x\in(f^{-1}[a-r,b+r])\backslash U\}>0.$$

**证** 直接由(PS)推出 $K_{ab}$紧. 若所要求 $r$ 不存在，则有 $x_n\in(f^{-1}[a-n^{-1},b+n^{-1}])\backslash U$；$|\mathrm{d}f(x_n)|\to0$. 由(PS)不妨设 $x_n\to x$，则显然 $x\in K_{ab}\backslash U$，得出矛盾. □

## §3 形变定理

本节中设 $f\in C^1(M)$满足(PS). 对任给$[a,b]\subset\mathrm{R}$，约定 $K_{ab}=K$

$\cap f^{-1}[a,b]$，$K_a=K\cap f^{-1}(a)$，$f_a=f^{-1}(-\infty,a]$，$K$ 是 $f$ 的临界点集. 这些记号在本章中将一直保持.

若 $A$ 是一拓扑空间，$B\subset A$，$h\in C(J\times A,A)$ 满足 $h_0=\mathrm{id}$，$h_1A\subset B$，$h_t|B=\mathrm{id}(\forall\, t\in J)$，则称 $h$ 为从 $A$ 到 $B$ 的强形变收缩，称 $B$ 为 $A$ 的强形变收缩核.

本节的主要结果指明，若 $K_{ab}=\varnothing$，则 $f_a$ 是 $f_b$ 的强形变收缩核，所需形变沿负伪梯度场的流线实现；进而可指出 $f_b$ 同胚于 $f_a$. 因此，一旦发现 $f_b$ 不同胚于 $f_a$，即可断定 $f^{-1}[a,b]$含 $f$ 的临界点. 临界点理论的基本思想乃源于此.

以下取定 $f$ 的伪梯度场 $\xi$，设 $\sigma$ 是 $-\xi$ 的流.

**7.3.1 引理** 设 $K_{ab}=\varnothing(a<b)$，则$\forall\, x\in f^{-1}[a,b]$，存在唯一 $\rho=\rho(a,x)\geqslant 0$，使 $f(\sigma(\rho,x))=a$；且 $x\mapsto\rho(a,x)$连续. 关于 $\rho(b,x)$有类似结论.

**注** $\rho(a,x)$可看作"从 $x$ 至 $f^{-1}(a)$的到达时间".

**证** 由 7.2.5，$\varepsilon\triangleq\inf\{|\mathrm{d}f(x)|: x\in f^{-1}[a,b]\}>0$. 取定 $x\in f^{-1}[a,b]$，令 $x(t)=\sigma(t,x)$，设$[0,\beta]$是 $x(t)$的极大正半定义区间. 若$x([0,\beta))\subset f^{-1}[a,b]$，则

$$\frac{\mathrm{d}}{\mathrm{d}t}f(x(t))\leqslant-|\mathrm{d}f(x(t))|^2\leqslant-\varepsilon^2 \tag{1}$$

(参考 §2(8))，于是

$$\beta\varepsilon^2\leqslant-\int_0^\beta\frac{\mathrm{d}}{\mathrm{d}t}f(x(t))\mathrm{d}t\leqslant b-a,$$

可见 $\beta<\infty$. 若 $0<\tau<t<\beta$，则结合(1)与 §1(3)，§2(2)有

$$\begin{aligned}\mathrm{d}(x(\tau),x(t))&\leqslant\int_\tau^t|x'(s)|\mathrm{d}s=\int_\tau^t|\xi(x(s))|\mathrm{d}s\\&\leqslant 2\int_\tau^t|\mathrm{d}f(x(s))|\mathrm{d}s\leqslant 2[(t-\tau)\int_\tau^t|\mathrm{d}f(x(s))|^2\mathrm{d}s]^{1/2}\\&\leqslant 2[(\tau-t)\int_\tau^t\frac{\mathrm{d}}{\mathrm{d}s}f(x(s))\mathrm{d}s]^{1/2}\leqslant 2\sqrt{(b-a)(t-\tau)},\end{aligned}\tag{2}$$

这推出 $x(\beta-0)$ 存在，与 $[0,\beta)$ 的极大性矛盾. 故必有 $\tau\in(0,\beta)$：$f(x(\tau))<a$；从而有 $\rho\in[0,\tau)$：$f(x(\rho))=a$. 由 §2(8) 知 $f(x(t))$ 严格下降，故如上的 $\rho$ 唯一. 对方程 $f(\sigma(\rho,x))=a$ 运用隐函数定理（因只需局部地运用隐函数定理，故可认为问题是在 Banach 空间中考虑），得出 $\rho(a,x)$ 对 $x$ 连续. □

以下设 $[a,b]\subset\mathbb{R}\,(a<b)$ 是取定的.

**7.3.2 定理** 设 $K_{ab}=\varnothing$，则存在从 $f_b$ 到 $f_a$ 的强形变收缩 $h$；当 $K\cap f^{-1}[a,\infty)=\varnothing$ 时 $f_b$ 可换成 $M$.

**证** 设 $\rho_a\triangleq\rho(a,x)$ 依 7.3.1，定义

$$h(t,x)=\begin{cases}x, & x\in f_a;\\ \sigma(\rho_a t,x), & x\in f_b\backslash f_a.\end{cases} \tag{3}$$

只需验证 $h:J\times f_b\to f_b$ 连续. 由 7.2.3 与 7.3.1 知 $h$ 限制在 $J\times f^{-1}[a,b]$ 上连续，因此只需证 $h$ 在任给点 $(t_0,x_0)\in J\times f^{-1}(a)$ 连续. 设 $(t,x)\in J\times(f_b\backslash f_a)$，则

$$\begin{aligned}&\mathrm{d}(h(t,x),h(t_0,x_0))\leqslant \mathrm{d}(x_0,x)+\mathrm{d}(x,h(t,x))\\ \leqslant\ &\mathrm{d}(x_0,x)+\int_0^{\rho t}|\xi(\sigma(s,x))|\mathrm{d}s\leqslant \mathrm{d}(x_0,x)+2\sqrt{\rho(b-a)},\end{aligned}$$

其中 $\rho=\rho(a,x)$，最后一步用到(2). 以上估计表明 $h$ 在 $(t_0,x_0)$ 连续.

若 $K\cap f^{-1}[a,\infty)=\varnothing$，则在(3)中以 $M$ 代 $f_b$，得到 $h:J\times M\to M$，以上论证依然有效. □

形变(3)构造简单且有明显的直观图景，但定义域受到限制. 今稍作修正以克服后一缺点.

**7.3.3 定理** 设 $K_{ab}=\varnothing$，则存在 $h\in C(J\times M,M)$，使 $h_0=\mathrm{id}$，$h$ 将 $f_b$ 强形变收缩到 $f_a$，且 $f\circ h_t\leqslant f$.

**证** 由 7.2.5，有 $c>b$：$k_{ac}=\varnothing$. 定义

$$h(t,x)=\begin{cases}\sigma(\rho_a t,x), & x\in f_b\backslash f_a;\\ \sigma\left(\dfrac{\rho_a-\rho_b)\rho_c t}{\rho_c-\rho_b},x\right), & x\in f_c\backslash f_b;\\ x, & x\in f_a\cup(M\backslash f_c),\end{cases} \tag{4}$$

其中 $\rho_a=\rho(a,x)$，$\rho_b$，$\rho_c$ 仿此. 类似于证 7.3.2，可验明 $h$ 合于定理所求. □

类似的方法可用来得出以下同胚结果.

**7.3.4 定理** 设 $K_{ab}=\varnothing$，$[c,d]\subset(a,b)$，则存在 $\varphi\in\text{Homeo}(M)$，使 $\varphi(f_d)=f_c$，$f\circ\varphi\leqslant f$，$\forall\, x\in f_a\cup(M\setminus f_b)$：$\varphi(x)=x$.

所需的 $\varphi$ 例如可定义如下（参照(4)）：

$$\varphi(x)=\begin{cases}\sigma\left(\dfrac{(\rho_c-\rho_d)\rho_a}{\rho_a-\rho_d},x\right), & x\in f_d\setminus f_a;\\ \sigma\left(\dfrac{(\rho_c-\rho_d)\rho_b}{\rho_b-\rho_b},x\right), & x\in f_b\setminus f_d;\\ x, & \text{其它}.\end{cases}\tag{5}$$

对于 $K_{ab}\neq\varnothing$ 的情况，只有在很特殊的条件下才有类似于 7.3.2 的结果，且所需形变有更复杂的构造.

**7.3.5 定理** 设 $f\in C^{2-0}(M)$（这意味着 $df\in C^{1-0}$），$f$ 在 $[a,b]$ 上仅有临界值 $a$，$K_a=\{x_1,x_2,\cdots,x_n\}$ 是有限集，则 $f_a$ 是 $f_b$ 的强形变收缩核.

**证** $\forall\, x\in f_b\setminus f_a$，以 $\sigma(\cdot,x)$ 记 IVP

$$\sigma'=((a-f(x))/\langle df(\sigma),\xi(\sigma)\rangle)\xi(\sigma),\sigma(0)=x\tag{6}$$

的解（因 $\xi,df\in C^{1-0}$，解必存在且唯一）. 直接验知

$$\frac{\mathrm{d}}{\mathrm{d}t}f(\sigma(t,x))=a-f(x),f(\sigma(t,x))=at+(1-t)f(x).\tag{7}$$

(7)表明 $f$ 沿 $\sigma(\cdot,x)$ 严格下降，$\forall\, t\in[0,1)$：$f(\sigma(t,x))>a$，因此 $\sigma(\cdot,x)$ 必在 $[0,1)$ 上定义且 $f(\sigma(t,x))\downarrow a(t\to 1)$.

令 $\mathrm{d}_i=\inf\limits_{0\leqslant t<1}d(\sigma(t,x),x_i)(1\leqslant i\leqslant n)$. 若 $\mathrm{d}_i>0(1\leqslant i\leqslant n)$，则类似于 7.3.1 的证法可得出 $\sigma(1-0,x)$ 存在. 若某个 $\mathrm{d}_i=0$，则必 $\sigma(1-0,x)=x_i$；否则，当 $t\to 1$ 时 $\sigma(t,x)$ 无限次相继穿过某对"球面"$S_k=\{y\in M:d(y,x_i)=k\varepsilon\}$，$k=1,2$，$\varepsilon>0$，而在 $S_1$ 与 $S_2$ 之间 $|(\mathrm{d}/\mathrm{d}t)\sigma(t,x)|$ 有界，此属不可能. 因此 $\sigma(1-0,x)$ 恒存在. 定义

$$h(t,x)=\begin{cases}\sigma(t,x),(t,x)\in[0,1)\times(f_b\setminus f_a);\\ \sigma(1-0,x),(t,x)\in\{1\}\times(f_b\setminus f_a);\\ x,\quad (t,x)\in J\times f_a.\end{cases}\tag{8}$$

只需验证 $h$ 在$\{1\}\times(f_b\setminus f_a)$与 $J\times f^{-1}(a)$上连续，不妨只考虑前者（后者是类似的）. 取定 $x\in f_b\setminus f_a$，不妨设 $h(1,x)=x_i\in K_a$，今证当 $t\uparrow 1,y\to x$ 时 $h(t,y)\to x_i$. $\forall\ \varepsilon>0$，取 $\delta_1,\delta_2>0$，使当 $d(x,y)<\delta_2<\varepsilon/2$时 $d(\sigma(1-\delta_1,y),x_i)<\varepsilon/2$. 可设 $\rho\triangleq\inf\{|df(y)|:\varepsilon/2\leqslant d(y,x_i)\leqslant\varepsilon\}>0$, $(b-a)\delta_1<\rho\varepsilon/4$. 今指明当 $d(x,y)<\delta_2, 1-\delta_1<t<1$ 时 $d(x_i,\sigma(t,y))<\varepsilon$. 否则，有 $y\in M, d(y,x)<\delta_2$，及$[\tau,t]\subset[1-\delta_1,1)$，使 $d(\sigma(\tau,y),x_i)=\varepsilon/2, d(\sigma(t,y),x_i)=\varepsilon, \forall\ s\in[\tau,t]:\varepsilon/2\leqslant d(\sigma(s,y),x_i)\leqslant\varepsilon$. 这推出

$$\frac{\varepsilon}{2}\leqslant \mathrm{d}(\sigma(\tau,y),\sigma(t,y))\leqslant\int_\tau^t\left|\frac{\mathrm{d}}{\mathrm{d}s}\sigma(s,y)\right|\mathrm{d}s\leqslant\frac{2}{\rho}(b-a)\delta_1<\frac{\varepsilon}{2},$$

得出矛盾. □

**7.3.6 定理** 设 $c\in\mathrm{R}$, $N$ 是 $K_C$ 的邻域，则存在 $\rho>r>0, h\in C(J\times M,M)$，使 $h_0=\mathrm{id}, h_1(f_{c+r}\setminus N^\circ)\subset f_{c-r}, h_t\in\mathrm{Homeo}(M)(\forall\ t\in J), \forall\ x\in M\setminus f^{-1}(c-\rho,c+\rho):h(t,x)=x$.

**证** 因 $K_C$ 紧(7.2.5)，可取 $\delta>0$，使$N_{6\delta}\triangleq\{x\leqslant M:\mathrm{d}(x,k_c)\leqslant 6\delta\}\subset N^\circ$. 由 7.2.5，有 $\rho>0$，使

$$\varepsilon\triangleq\inf\{|\mathrm{d}f(x)|:x\in(f^{-1}[c-\rho,c+\rho])\setminus N_\delta\}>0.$$

令 $\varphi(x)=\mathrm{d}(x,N_\delta)[\mathrm{d}(x,N_\delta)+\mathrm{d}(x,M\setminus N_{2\delta})]^{-1}, x\in M,$ (9)

则 $\varphi\in C^{1-0}(M,J), \varphi|N_\delta=0, \varphi|(M\setminus N_{2\delta})=1$. 取 $0<r<\min\{\rho,\varepsilon\delta\varepsilon/8\}$，类似于(9)定义 $\psi\in C^{1-0}(M,J)$，使 $x\in M\setminus f^{-1}[c-\rho,c+\rho]\Rightarrow\psi(x)=0, x\in f^{-1}[c-r,c+r]\Rightarrow\psi(x)=1$. 定义

$$\eta(x)=-\varphi(x)\psi(x)|\xi(x)|^{-1}\xi(x),x\in\tilde{M};\tag{10}$$

当 $x\in K$时令 $\eta(x)=0$. 则 $\eta$ 是 $M$ 上的 $C^{1-0}$向量场，$|\eta(x)|\leqslant 1$. 设是 $\eta$ 决定的流(7.2.3)，今证 $h\triangleq\theta|J\times M$ 即合于定理要求. 只需验证

$\eta_1(f_{c+r}\setminus N_{6\delta})\subset f_{c-r}$. 取定 $\bar{x}\in f_{c+r}\setminus N_{6\delta}$，令 $x(t)=h(t,\bar{x})$. 若 $x(1)\bar{\in} f_{c-r}$，则 $x(J)\subset f^{-1}(c-r,c+r]$. 令 $t_0=4r/\varepsilon$，则 $0<t_0<1$. 若 $x([0,t_0))\cap N_{2\delta}=\varnothing$，则

$$2r>f(\bar{x})-f(x(t_0))=-\int_0^{t_0}\langle \mathrm{d}f(x(t)),\eta(x(t))\rangle \mathrm{d}t$$

$$=\int_0^{t_0}\frac{\langle \mathrm{d}f(x(t)),\xi(x(t))\rangle}{|\xi(x(t))|}\mathrm{d}t\geqslant\frac{\varepsilon t_0}{2}=2r,$$

得出矛盾. 故∃ $t_1\in(0,t_0)$: $x(t_1)\in N_{2\delta}$, $x([0,t_1))\cap N_{2\delta}=\varnothing$;

$$\mathrm{d}(x(t_1),\bar{x})\leqslant\int_0^{t_1}|x'(t)|\mathrm{d}t=\int_0^{t_1}|\eta(x(t))|\mathrm{d}t$$

$$\leqslant t_1\leqslant 2\varepsilon^{-1}[f(\bar{x})-f(x(t_1))]\leqslant 4r\varepsilon^{-1}\leqslant 4\delta,$$

这推出 $\mathrm{d}(\bar{x},K_c)\leqslant\mathrm{d}(\bar{x},x(t_1))+2\delta\leqslant 6\delta$，与 $\bar{x}\bar{\in} N_{6\delta}$ 矛盾. 因此 $x(1)\in f_{c-r}$，如所要证. □

## §4 Minimax 定理

依然设 $f\in C^1(M)$ 满足(PS).

在很多情况下，临界点具有"鞍点"特征. 最简单的例子是 $\varphi(x,y)=x^2-y^2(x,y\in\mathbb{R})$，对于它有

$$\varphi(0,0)=\inf_{r>0}\sup_{|x|=r,|y|<\infty}\varphi(x,y)=0,$$

临界值 0 被表成"minimax 值". 这启示出一个在本章有基本意义的问题：对何种集族 $\sigma\subset 2^M$,

$$c_\sigma\triangleq\inf_{S\in\sigma}\sup_{x\in S}f(x)\tag{1}$$

是 $f$ 的临界值? 对此问题的解答构成一系列"minimax 定理". 基本的结论是：若 $\sigma\subset 2^M$ 对一定映射族 $\Phi$ 不变(即∀ $\varphi\in\Phi, S\in\sigma$: $\varphi S\in\sigma$)，则 $c_\sigma$ 是临界值. 下面是 $\Phi$ 的两种可行的选择：

$$\Phi_a=\{\varphi: \exists\ b>a \text{ 及如 7.3.3 之 } h \text{ 使 } \varphi=h_1\};\tag{2}$$

$$\Phi_{ab}=\{\varphi: \exists\ [c,d]\subset(a,b) \text{ 使 } \varphi \text{ 如 7.3.4}\}.\tag{3}$$

(2)(3)皆与 $f$ 有关，必要时分别写作 $\Phi_a(f)$ 与 $\Phi_{ab}(f)$.

**7.4.1Minimax 原理**　设 $\sigma\subset 2^M$，$c=c_\sigma$(依(1))有限，当 $r>0$ 充分小时 $\sigma$ 对 $\Phi_{c-r}$ 或 $\Phi_{c-r,c+r}$ 不变，则 $c$ 是 $f$ 的临界值.

**证**　要证者 $K_c\neq\varnothing$. 若 $K_c=\varnothing$，则对充分小的 $r>0$ 有 $K_{c-r,c+r}=\varnothing$(7.2.5). 取 $S\in\sigma$，使 $\sup f(S)<c+r/2$. 由(2)与 7.3.3 有 $\varphi\in\Phi_{c-r}$，使 $\varphi(f_{c+r})\subset f_{c-r}$. 若 $\sigma$ 对 $\Phi_{c-r}$ 不变，则 $\varphi S\in\sigma$，于是由(1)有

$$c\leqslant\sup f(\varphi S)\leqslant\sup f(\varphi(f_{c+r}))\leqslant\sup f(f_{c-r})\leqslant c-r,$$

得出矛盾. 其次，由(3)与 7.3.4 有 $\varphi\in\Phi_{c-r,c+r}$，使 $\varphi(f_{c+r/2})=f_{c-r/2}$. 若 $\sigma$ 对 $\Phi_{c-r,c+r}$ 不变，则 $\varphi S\in\sigma$，

$$c\leqslant\sup f(\varphi(f_{c+r/2}))=\sup f(f_{c-r/2})\leqslant c-\frac{r}{2},$$

亦得矛盾. 因此必 $K_c\neq\varnothing$. □

7.4.1 的结论远不是平凡的，而其表述与证明却极其简单，以至令人怀疑它是否有实质性的应用. 确实，倘无有效方法构成定理所要求的集族 $\sigma$，定理将无从利用. 幸而对 $\sigma$ 已有某些成功的选择. 最简单的选择莫过于取 $\sigma=\{\{x\}:x\in M\}$ 或 $\sigma=\{M\}$，这样的 $\sigma$ 对任何 $\Phi_{ab}$ $(a<b)$ 不变，而 $c_\sigma=\inf f(M)$ 或 $C_\sigma=\sup f(M)$，于是由 7.4.1 有

**7.4.2 推论**　若 $c=\inf f(M)>-\infty$ [$c=\sup f(M)<\infty$]，则 $f$ 有极小[极大]点 $\bar{x}$，且 $\bar{x}\in K_c$.

这是一个虽非平凡但较简单的结果. 更深入的结论自然有赖于对 $\sigma$ 的更精巧的选择，使其既不失之过宽(如 7.4.1)，亦不失之过窄(如 7.4.2).

迄今广为采用的一种构成集族 $\sigma$ 的方法基于"环绕"概念. 给定 $A,B\subset M$ 及 $L\subset A$，若 $L\cap B=\varnothing$，$\forall\ \varphi\in C(A,M):\varphi|L=\mathrm{id}\Rightarrow B\cap\varphi A\neq\varnothing$，则说 $B$ 穿过 $A$ 环绕 $L$，或简单地说"$B$ 环绕 $L$". 直观上可设想 $A$ 是以 $L$ 为"边界"的"区域"，$\varphi A$ 则是以 $L$ 为边界的"曲面"，$B$ 是恒穿过 $\varphi A$ 内部的一条"曲线". 令

$$\Phi=\{\varphi\in C(A,M):\varphi|L=\mathrm{id}\},\tag{4}$$

$\sigma=\{\varphi A:\varphi\in\Phi\}$，$c_\sigma$ 依(1)，则 $\sup f(L)\leqslant c_\sigma\leqslant\sup f(A)$；当 $B$ 环绕 $L$ 时 $\inf f(B)\leqslant c_\sigma$.

**7.4.3 环绕定理** 设 $B$ 环绕 $L$，$f(L)\leqslant\alpha<\beta\leqslant f(B)$，$\sup f(A)<\infty$，则 $f$ 有临界值 $c\geqslant\beta$.

**证** 设 $\Phi$ 依(4)，$\sigma=\{\varphi A:\varphi\in\Phi\}$，$c=c_\sigma$，则 $\alpha<\beta\leqslant c<\infty$. 任给 $r\in(0,c-\alpha)$，今证 $\sigma$ 对 $\Psi\triangleq\Phi_{c-r}$不变(从而由 7.4.1 知 $c$ 是临界值).

任给 $\varphi\in\Phi$，$\psi\in\Phi$，$g\triangleq\psi\circ\varphi\in C(A,M)$；$L\subset f_\alpha\subset f_{c-r}\Rightarrow g|L=\Psi|L=\mathrm{id}$，可见 $g\in\Phi$，从而 $\psi(\varphi A)=g(A)\in\sigma$，这表明 $\sigma$ 对 $\Psi$ 不变. □

定理中的 $A,B,L$ 仍有很大的一般性，致使 7.4.3 有较大的应用余地. 现在的问题是如何适当地选择"环绕"，以使 7.4.3 具体化. 最简单的选择莫过于取 $A$ 为一线段而 $L$ 是其两端点，这导致著名的"山路引理"(Ambrosetti-Rabinowitz[4]).

**7.4.4 山路引理** 设 $f\in C^1(X)$满足(PS)，$\Omega$ 是 $X$ 的 0-邻域，$x_0\in X\backslash\overline{\Omega}$，$\max\{f(0),f(x_0)\}<\beta\leqslant f(\partial\Omega)$，则 $f$ 有临界值 $c\geqslant\beta$，从而方程 $f'(x)=0$ 有非零解.

**证** 对 $M=X$，$A=[0,x_0]$，$L=\{0,x_0\}$，$B=\partial\Omega$ 应用 7.4.3 即得. □

**注** 对如上的 $A,L$，当 $\varphi\in C(A,X)$，$\varphi|L=\mathrm{id}$ 时，$\varphi A$ 似乎是一条越过"山岭"$\partial\Omega$ 的"山路"；$f$ 的临界值 $c=\inf\limits_{\varphi}\sup f(\varphi A)$似乎是在穿过某个"山坳"的一条山路的顶点上达到. "山路引理"之名盖出于此.

山路引理对于微分方程问题有深刻的应用. 鉴于它的重要性，有不少工作致力于寻求它的多种推广与变种([145,146,203]). 山路引理简单而又深刻，在现代分析定理中堪为典范. 但它远未穷尽环绕定理的潜力. 实际上，几何上自然且分析上可用的环绕甚多，其中每种都可能用来得出 7.4.3 的某种具体化，它们或多或少具有与山路引理相当的作用. 下面考虑两个例子.

**7.4.5 命题** 设 $X=X_1\oplus X_2$，$\dim X_1<\infty$，$r>0$，则 $X_2$ 穿过 $A=$

$X_1\cap\overline{B}_r(0)$与 $L=X_1\cap\partial B_r(0)$环绕.

证　设 $P:X\to X_1$ 是投影. 给定 $\varphi\in C(A,X)$, $\varphi|L=\mathrm{id}$, 则 $X_2\cap\varphi A\neq\varnothing\Leftrightarrow P\varphi(x)=0$ 在 $A$ 内有解. 因 $P\circ\varphi|L=\mathrm{id}$, 故所要结论由度理论得出. □

**7.4.6 命题**　设 $X=X_1\oplus X_2$, $\dim X_1<\infty$, $e\in X_2$, $|e|=1$, $A=(X_1\cap B_r(0))+[0,\delta]e$, $L$ 是 $A$ 在 $X_1\oplus\mathrm{Re}$ 中的边界, $B=X_2\cap\partial B_\rho(0)$, $r>0$, $0<\rho<\delta$, 则 $B$ 穿过 $A$ 环绕 $L$.

证　设 $P:X\to X_1$ 与 $Q:X\to X_2$ 是投影, $Y=X_1\oplus\mathrm{R}e$. 给定 $\varphi\in C(A,X)$, $\varphi|L=\mathrm{id}$, 今证 $B\cap\varphi A\neq\varnothing$, 这相当于证 $\exists\ x\in A$: $P\varphi(x)+(|Q\varphi(x)|-\rho)e=0$. 证法在于巧妙地构造同伦并运用度论证. 令

$$h(t,x)=P(t'x+t\varphi(x))+t'Qx+(t|Q\varphi(x)|-\rho)e,$$

其中 $t'=1-t$, 则 $h\in C(J\times A,Y)$; $\forall\ x\in L$:

$$h(t,x)=Px+t'Qx+(t|Qx|-\rho)e\neq 0$$

(否则易推出 $x=\rho e$, 与 $x\in L$ 矛盾); $h(0,x)=x-\rho e$, $h(1,x)=P\varphi(x)+(|Q\varphi(x)|-\rho)e$. 于是

$$d(h_1,A,0)=d(h_0,A,0)=d(I,A,\rho e)=1,$$

这得出 $B\cap\varphi A\neq\varnothing$. □

注　7.4.6 中的 $A$ 可换成 $Y$ 中的"半球"(见[204])

$$A=\{x_1+\lambda e;x_1\in X_1,\lambda\in\mathrm{R}_+,|x_1|^2+\lambda^2\leqslant r\}.$$

对于以上考虑的几种情况, 环绕的直观含义是明显的.

下面考虑 7.4.3 的两个变种.

**7.4.7 对偶环绕定理**　设 7.4.3 之条件满足, $\Phi^*=\{\varphi\in\mathrm{Homeo}(M):\varphi|L=\mathrm{id}\}$, $\sigma^*=\{\varphi B:\varphi\in\Phi^*\}$, $c^*=\sup\limits_{S\in\sigma^*}\inf f(S)$, $c$ 依 7.4.3, 则 $c^*\leqslant c$ 且 $c^*$ 是 $f$ 的临界值.

证　$\forall\ \psi\in\Phi^*$, $\varphi\in\Phi$, 有 $\psi^{-1}\varphi\in\Phi$, 于是 $\psi B\cap\varphi A=\psi(B\cap\psi^{-1}\varphi A)\neq\varnothing$, 这推出 $c^*\leqslant c$. 因 $-c^*=\inf\limits_{S\in\sigma^*}\sup -f(S)$, 故只要证 $\sigma^*$ 对 $\Psi\triangleq\Phi_{-c^*-r,-c^*+r}(-f)$不变, $r\in(0,\beta-\alpha)$. $\forall\ \varphi\in\Phi^*$, $\psi\in\Psi$, $g\triangleq\psi\circ\varphi\in$

Homeo$(M)$;id$\in\Phi^*\Rightarrow c^*\geqslant\beta$，于是 $L\subset f_a\subset f_{c^*-r}$，这推出 $g|L=\psi|L=\mathrm{id}$，$g\in\Phi^*$，因此 $\sigma^*$ 对 $\Psi$ 不变. □

**7.4.8 定理**（张恭庆[204]） 设 $A,B,L$ 如 7.4.3，$f\in C^{2-0}(M)$ 满足(PS)，$f(L)\leqslant\alpha<f(B)$，$\sup f(A)<\infty$，$c$ 依 7.4.3，则 $c$ 是 $f$ 的临界值，$c>\alpha$，或 $c=\alpha$ 且 $|K_c|=\infty$.

**证** 显然 $c\geqslant\alpha$. 若 $c>\alpha$，则 7.4.3 的证法依然可用. 下面设 $c=\alpha$，可设$\exists\ r>0$：$L\cap f^{-1}(c,c+r]=\varnothing$. 若 $|K_c|<\infty$，则由 7.3.5 有 $\psi\in C(f_{c+r},f_c)$：$\psi|f_c=\mathrm{id}$. 设 $\Phi$ 依(4)，取 $\varphi\in\Phi$，使 $\varphi A\subset f_{c+r}$. 因 $g\triangleq\psi\circ\varphi\in C(A,M)$，$g|L=\psi|L=\mathrm{id}$，故 $g\in\Phi$，于是有 $x\in B\cap g(A)$. 这推出

$$\alpha<f(x)\leqslant\sup f(g(A))\leqslant\sup f(\psi(f_{c+r}))\leqslant c=\alpha,$$

得出矛盾. 因此 $|K_c|=\infty$. □

## §5 一个算子方程

山路引理一类的结果对于方程问题有多种多样的应用，且已产生丰富的结果，要作任何形式的简短概括都是不可能的. 本节所举的例子取自[204;Ch.3，§3]，在方法上它颇具典型性，因此用作说明是适当的.

设$(\Omega,\mu)$是一测度空间，$0<\mu\Omega<\infty$，记 $L^p=L^p(\Omega,\mathrm{R}^n)$. 给定自伴算子 $A:D(A)\subset L^2\to L^2$，$f\in C^1(\mathrm{R}^n)$，$Fu(x)=f'(u(x))$，考虑半线性算子方程

$$Au+Fu=0. \qquad (1)$$

取定 $p\in(2,\infty)$，$p^{-1}+q^{-1}=1$. 若 $u\in L^p$ 满足

$$\langle u,Av\rangle+\langle v,Fu\rangle=0(v\in D(A)\cap L^p), \qquad (2)$$

其中$\langle\cdot,\cdot\rangle$表 $L^p$ 与 $L^q$ 之间的配对，则称 $u$ 为(1)的弱解. 方程(1)是多种具体方程的抽象形式，因此有很大的重要性. 今用山路引理建立：

**7.5.1 定理**　设以下条件满足：

$(H_1)A$ 有负特征值 $\lambda$ 且 $R(A)\subset L^2$ 闭；

$(H_2)A^{-1}P$ 可扩张为紧线性算子 $K:L^q\to L^p$，此处 $P:L^2\to R(A)$ 是正投影；

$(H_3)f\in C^1(\mathrm{R}^n)$ 严格凸，$f(0)=0$，$f(z)=o(|z|^2)(z\to 0)$；

$(H_4)\exists\ \rho,C>0:|z|\geqslant\rho\Rightarrow f(z)\leqslant\min\{C|z|^p,p^{-1}f'(z)\cdot z\}$，则方程(1)有非零弱解.

**证**　令 $g=f^*$，由下面就要指明的不等式(4)有 $|f'(z)|\to\infty$ $(|z|\to\infty)$，于是用 5.8.4 与 4.1.5 得 $g\in C^1(\mathrm{R}^n)$ 严格凸，$g'=f'^{-1}$. 由$(H_3)$有 $f'(0)=0$，从而 $g'(0)=0$；再由 5.8.3(i)得 $g(0)=0$. $z=0$ 是 $f,g$ 的唯一极小点(4.1.10).

1°首先对 $f,g$ 建立以下不等式：

$$\mathrm{const}(|z|^p-1)\leqslant f(z)\leqslant\mathrm{const}(|z|^p+1);\tag{3}$$

$$\mathrm{const}(|z|^{p-1}-1)\leqslant|f'(z)|\leqslant\mathrm{const}(|z|^{p-1}+1);\tag{4}$$

$$\mathrm{const}(|u|^q-1)\leqslant g(u)\leqslant\mathrm{const}(|u|^q+1);\tag{5}$$

$$\mathrm{const}(|u|^{q-1}-1)\leqslant|g'(u)|\leqslant\mathrm{const}(|u|^{q-1}+1);\tag{6}$$

$$g(u)\geqslant\begin{cases}C_\delta|u|^2, & |u|\leqslant\delta;\\ C'_\delta|u|^q, & |u|\geqslant\delta,\end{cases}\tag{7}$$

其中 const 表某个正常数，其具体数值依情况而定，不必准确知道(下文仿此)；(7)中的 $\delta>0$ 是任给的，$C_\delta,C'_\delta>0$ 依赖于 $\delta$，当 $\delta\downarrow 0$ 时 $C_\delta\to\infty$. (3)的后一不等号直接由条件$(H_4)$推出. 若 $|z|\geqslant\rho,t\geqslant 1$，则由$(H_4)$有 $pf(tz)\leqslant f'(tz)\cdot tz$，于是

$$\ln t^p=\int_1^t\frac{p\mathrm{d}s}{s}\leqslant\int_1^t\frac{f'(sz)\cdot z}{f(sz)}\mathrm{d}s$$

$$=\int_1^t\frac{\mathrm{d}}{\mathrm{d}s}\ln f(sz)\mathrm{d}s=\ln\frac{f(tz)}{f(z)},$$

故得 $f(tz)\geqslant t^pf(z)(t\geqslant 1,|z|\geqslant\rho)$. 由此推出

$$f(z)\geqslant(|z|/\rho)^pf(\rho z/|z|)\geqslant\mathrm{const}|z|^p(|z|\geqslant\rho),$$

这得出(3). (3)结合($H_4$)得出$|f'(z)|\geqslant \text{const}(|z|^{p-1}-1)$. 若$|z|\geqslant 1$,则用(3)与 4.1.5 得

$$\begin{aligned}|f'(z)| &= \sup_{|u|=|z|}\frac{1}{|z|}|f'(z)\cdot u| \\ &\leqslant \sup_{|u|=|z|}\frac{1}{|z|}[|f(z+u)|+|f(z)|] \\ &\leqslant \sup_{|u|=|z|}\text{const}\frac{1}{|z|}(|z+u|^p+|z|^p+2) \\ &\leqslant \text{const}(|z|^{p-1}+1).\end{aligned}$$

$|z|\geqslant 1$ 的限制显然可以除去,于是(4)得证. 利用 $g'=f'^{-1}$从(4)推出(6). 设(3)式两端之常数分别为 $c,c_1$,取 $b>0$ 使 $b-c_1b^p>0$,则

$$g(u)\leqslant \sup_z(|u||z|-C|z|^p+C)\leqslant \text{const}(|u|^q+1);$$

$$\begin{aligned}g(u) &\geqslant \sup\{u\cdot z-f(z):|z|=b|u|^{q-1}\} \\ &\geqslant (b-c_1b^p)|u|^q-c_1,\end{aligned}$$

这得出(5). $\forall\,\delta>0$,令 $C_\delta=\inf_{|u|\leqslant\delta}|u|^{-2}g(u)$,$C'_\delta=\inf_{|u|\geqslant\delta}|u|^{-q}g(u)$,则 $C_\delta,C'_\delta$ 使(7)成立,且 $0\leqslant C_\delta<\infty$,$0<C'_\delta<\infty$. $\forall\,\epsilon>0$,由($H_3$)有 $\delta>0$: $|z|\leqslant\delta\Rightarrow|f(z)|\leqslant\epsilon|z|^2$;这推出$|u|\leqslant 2\epsilon\delta$ 时

$$g(u)\geqslant \sup_{|z|=|u|/2\epsilon}[u\cdot z-f(z)]\geqslant\frac{1}{4\epsilon}|u|^2,$$

故

$$\lim_{|u|\to 0}|u|^{-2}g(u)=\infty. \tag{8}$$

由(8)推出 $C_\delta>0$ 且 $C_\delta\to\infty(\delta\downarrow 0)$.

令 $Gv(x)=g'(v(x))$,则 $G=F^{-1}$;(4)推出 $F\in C(L^p,L^q)$,(6)推出 $G\in C(L^q,L^p)$(参看[204;p.5]).

2°转化方程. 令

$$X=\{v\in L^q:\langle u,v\rangle=0(\forall\, u\in N(A)\cap L^p)\},$$

则 $X$ 是 $L^q$ 的闭子空间,若$(v,w)\in X\times(N(A)\cap L^p)$满足

$$Kv+Gv=w, \tag{9}$$

则 $u=Gv$ 必为(1)的弱解. 事实上,$\forall\, z\in D(A)\cap L^p$,有

$$\langle u, Az\rangle + \langle z, Fu\rangle = \langle w - Kv, Az\rangle + \langle z, v\rangle$$
$$= -\langle Kv, Az\rangle + \langle z, v\rangle = \langle z, v - AKv\rangle = 0,$$

即(2)满足. 注意 $v\neq 0\Rightarrow u\neq 0$. 定义

$$\mathcal{J}(v) = \frac{1}{2}\langle Kv, v\rangle + \int_\Omega g(v), v \in L^q, \tag{10}$$

此处及以下皆省去积分中的 $dx$. 由(10)直接看出 $\mathcal{J}\in C^1(L^q)$, $\mathcal{J}'(v) = Kv + Gv$. 令 $J = \mathcal{J}|X$, 则 $J\in C^1(X)$, $J'(v) - \mathcal{J}'(v)\in X^\perp = N(A)\cap L^p(v\in X)$. 若 $J'(v)=0$, 则有 $w\in N(A)\cap L^p$ 使 $\mathcal{J}'(v) = w$, 于是 $(v,w)$ 满足(9). 因此问题归于证 $J$ 有非零临界点.

3° 证 $J$ 满足(PS). 设 $\{v_k\}\subset X$, $J(v_k)$ 有界, $w_k \triangleq J'(v_k)\to 0(L^p)$, 设 $w_k = Kv_k + Gv_k + z_k$, $z_k\in X^\perp$. 首先证 $v_k$ 有界. 若 $u = f'(z)$, 则由 $(H_4)$, (3), (6)及 5.8.3(i)有

$$g(u) - \frac{1}{2}g'(u)\cdot u = u\cdot z - f(z) - \frac{1}{2}f'(z)\cdot z$$
$$= \frac{1}{2}f'(z)\cdot z - f(z) \geqslant \text{const}(|z|^p - 1)$$
$$= \text{const}[|g'(u)|^p - 1] \geqslant \text{const}(|u|^q - 1),$$

于是 $\text{const}|u|^q \leqslant g(u) - 2^{-1}g'(u)\cdot u + \text{const}$. 以 $u = v_k$ 代入得

$$\text{const}|v_k|_q^q \leqslant \int_\Omega g(v_k) - \frac{1}{2}\langle Gv_k, v_k\rangle + \text{const}$$
$$= J(v_k) - \frac{1}{2}\langle Kv_k + Gv_k, v_k\rangle + \text{const}$$
$$= J(v_k) - \frac{1}{2}\langle w_k, v_k\rangle + \text{const},$$

由此已可看出 $v_k$ 有界. 不妨设 $v_k \rightharpoonup v\in X$. 因 $K$ 紧, 故 $Kv_k\to Kv(L^p)$. 因 $g$ 凸, 由 4.1.5 有

$$\int_\Omega g(v) \leqslant \varliminf_k \int_\Omega [g(v_k) - g'(v)\cdot(v_k - v)]$$
$$= \varliminf_k \int_\Omega g(v_k) \leqslant \varlimsup_k \int_\Omega g(v_k)$$

$$\leqslant \overline{\lim_k}\int_\Omega [g(v) - g'(v_k)\cdot(v - v_k)]$$

$$= \int_\Omega g(v) + \overline{\lim_k}\langle Gv_k, v_k - v\rangle$$

$$= \int_\Omega g(v) + \overline{\lim_k}\langle w_k - Kv_k, v_k - v\rangle = \int_\Omega g(v),$$

故
$$\overline{\lim_k}\int_\Omega g(v_k) = \int_\Omega g(v). \tag{11}$$

(11)蕴涵 $v_k \to v(L^q)$，其证明留到最后一步.

4°对 $J$ 用山路引理. 显然 $J(0)=0$. 由($H_1$)有 $u\in L^2$：$Au=\lambda u$；易验知 $u\in X$. 因 $\lambda<0$，故 $t\to\infty$ 时

$$J(tu) = \frac{t^2}{2\lambda}|u|_2^2 + \int_\Omega g(tu)$$

$\leqslant \frac{t^2}{2\lambda}|u|_2^2 + \text{const}(t^q|u|_q^q + 1) \to -\infty$. 于是只要证∃ $r,m>0$：$J(\partial B_r(0))\geqslant m$. 取充分小的 $\delta, r>0$，设 $v\in X$，$|v|_q = r$，$k$ 记 $K$ 的算子范数，则

$$J(v) \geqslant \int_\Omega g(v) - \frac{1}{2}|\langle Kv, v\rangle|$$

$$\geqslant C_\delta \int_{|v|\leqslant\delta} |v|^2 + C'_\delta \int_{|v|\geqslant\delta} |v|^q - \frac{k}{2}|v|_q^2$$

$$\geqslant \left[C_\delta \int_{|v|\leqslant\delta} |v|^2 - 2k\left(\int_{|v|\leqslant\delta} |v|^q\right)^{2/q}\right]$$

$$+ \left[C'_\delta \int_{|v|\geqslant\delta} |v|^q - 2k\left(\int_{|v|\leqslant\delta} |v|^q\right)^{2/q}\right]$$

$$\geqslant \text{const}\left[\left(\int_{|v|\leqslant\delta} |v|^q\right)^{2/q} + \left(\int_{|v|\geqslant\delta} |v|^q\right)^{2/q}\right] \geqslant \text{const}\,|v|_q^2$$

其中用到初等不等式 $a^\alpha + b^\alpha \leqslant (a+b)^\alpha \leqslant 2^\alpha(a^\alpha + b^\alpha)$ $(a,b>0, \alpha>1)$. 因此 $J$ 有非零临界点.

5°证 $v_k\to v(L^q)$，$v_k, v$ 依证明第 3°段. ∀ $u\in L^q$，依(7)与 Hölder 不等式有

$$\int_{\Omega} g(u) \geqslant C_\delta \int_{|u|\leqslant\delta} |u|^2 + C'_\delta \int_{|u|>\delta} |u|^q$$

$$\geqslant C_\delta(\mu\Omega)^{1-2/q}\left(\int_{|u|\leqslant\delta} |u|^q\right)^{2/q} + C'_\delta \int_{|u|>\delta} |u|^q.$$

在所得不等式中以 $v_k-v$ 代 $u$ 看出，$v_k \to v(L^q)$ 蕴涵于

$$\lim_k \int_{\Omega} g(v_k - v) = 0. \tag{12}$$

设 $a>\delta>0$，$\Omega_1=\Omega(|v|>a)$，$\Omega_{2k}=\Omega(|v|\leqslant a)\cap\Omega(|v_k-v|\leqslant\delta)$，$\Omega_{3k}=\Omega\backslash(\Omega_1\cup\Omega_{2k})$. 由(5)易推出

$$g(2u) \leqslant \mathrm{const}[g(u)+1]. \tag{13}$$

利用 $g$ 的凸性及(13)得

$$I_1 \triangleq \int_{\Omega_1} g(v_k - v) \leqslant \frac{1}{2}\int_{\Omega_1}[g(2v_k) + g(-2v)]$$

$$\leqslant \mathrm{const}\int_{\Omega_1}[g(v_k) + g(-v) + 1]. \tag{14}$$

由 $v_k \rightharpoonup v$ 与(11)不难推出，对任何 $\mu$-可测集 $S\subset\Omega$；有 $\int_S g(v_k) \to \int_S g(v)$；由此推出 $g(v_k)$在 $\Omega$ 上有等度绝对连续积分. $\forall\ \varepsilon>0$，取 $a$ 充分大(从而 $\mu\Omega_1$ 充分小)，使 $I_1<\varepsilon$. 固定 $a$，取 $\delta$ 充分小，使

$$I_2 \triangleq \int_{\Omega_{2k}} g(v_k - v) < \varepsilon.$$

固定 $\delta$，由 $g$ 严格凸用 4.1.5 不难证得

$$m \triangleq \inf_{|z|\leqslant a,\,|u-z|\geqslant\delta}[g(u) - g(z) - g'(z)\cdot(u-z)] > 0.$$

由

$$m\mu\Omega_{3k} \leqslant \int_{\Omega_{3k}}[g(v_k) - g(v) - g'(v)\cdot(v_k - v)]$$

$$\leqslant \int_{\Omega}[g(v_k) - g(v) - g'(v)\cdot(v_k - v)]$$

$$= \int_{\Omega}[g(v_k) - g(v)] - \langle Gv, v_k - v\rangle \to 0,$$

推出 $\mu\Omega_{3k}\to 0$，于是用一类似于(14)的式子得出

$$I_3 \triangleq \int_{\Omega_{3k}} g(v_k - v) \to 0.$$

因 $\int_\Omega g(v_k-v)=I_1+I_2+I_3$，综合以上所证得出(12). □

**7.5.2 例** 设 $H\in C^1(\mathrm{R}^n\times\mathrm{R}^n)$，考虑 Hamilton 系统

$$x' = -H_p(x,p), p' = H_x(x,p). \tag{15}$$

(15)可缩写成 $z'=JH'(z), z=(x,p)^T, J=\begin{pmatrix}0 & -I\\ I & 0\end{pmatrix}$. 求(15)的 $2\pi$-周期解相当于求 $z:S^1\to\mathrm{R}^{2n}$，使之满足

$$Az + H'(z) = 0, \tag{16}$$

其中 $A=J(d/dt)$. 可验证 $A$ 在 $L^2(S^1,\mathrm{R}^{2n})$上有自伴扩张且满足条件$(\mathrm{H}_1)(\mathrm{H}_2)$(依 7.5.1)(参考[204;ch. 1])，于是当 $H$ 满足 7.5.1 之条件$(\mathrm{H}_3)(\mathrm{H}_4)$时(15)有非零 $2\pi$-周期解.

## §6 拓扑指标

§4 中给出的 minimax 定理仅能断定临界点存在，不能提供关于临界点个数的信息，因而被称为"弱 minimax 定理". 一类源于 Lusternik-Schnirelman 的拓扑方法的"强 minimax 定理"能给出临界点个数的某种估计，因而成为临界点理论中一种强有力的工具. Lusternik-Schnirelman 理论的基本思想是：构成有一定性质的"拓扑指标"$i$，依 $i(A)$之大小对 $M$ 的一定子集进行分类，形成一递降的集族序列 $\sigma_1\supset\sigma_2\supset\cdots$，设 $c_n=c_{\sigma_n}$依 §4(1)，则$\{c_n\}$是一增序列，它往往包含较多的临界点信息. 选择不同的拓扑指标，就得到形式多样的"Lusternik-Schnirelman 型定理".

拓扑指标概念通常与一定的连续群作用联系在一起. 设 $G$ 是一紧拓扑(乘法)群，$e$ 记其群幺元. 若 $T\in C(G\times X,X)$满足 $T_{ab}=T_aT_b, T_e=I$，此处 $T_a=T(a,\cdot)(a,b\in G)$，则称 $T$ 为 $G$ 对 $X$ 的连续

群作用. 以下设 $T:G\times X\to X$ 是给定的连续群作用, 且假定 $T_a(a\in G)$ 恒为等距同构. 约定 $a\cdot x=T_ax$, $A\cdot V=\{a\cdot x:a\in A,x\in V\}(A\subset G,V\subset X)$. 设 $V\subset X$, $f:V\to\mathrm{R}$, $\varphi:V\to X$. 若 $\forall\ a\in G:a\cdot V=V$, 则说 $V$ 是 $T(G)$-不变集. 若 $V$ 为 $T(G)$-不变, $\forall\ a\in G:f\circ T_a=f[\varphi\circ T_a=T_a\circ\varphi]$, 则说 $f$ 为 $T(G)$-不变[$\varphi$ 为 $T(G)$-等变]. 当 $G=\{e\}$ 时群作用是平凡的. 最简单的非平凡群作用是 $T(Z_2)$ 作用: $Z_2=\{0,1\}$, $T_0=I$, $T_1=-I$; 相对于 $T(Z_2)$ 的不变集、不变泛函与等变映射分别是对称集、偶泛函与奇映射.

以下设 $M$ 是 $X$ 的 $C^r$ 子流形, 以 $\sigma_G$ 记 $M$ 的 $T(G)$-不变闭子集之全体, 假定 $M\in\sigma_G$. 若 $G=\{e\}$, 则不必假定 $M$ 是 $X$ 的子流形.

**7.6.1 定义** 称集函数 $i:\sigma_G\to Z_+\cup\{\infty\}$ 为 $M$ 上的 $T(G)$-指标, 若它有性质: (i) 单调性: $A\subset B\Rightarrow i(A)\leqslant i(B)$; (ii) 次可加性: $i(A\cup B)\leqslant i(A)+i(B)$; (iii) 超变性: 若 $h\in C(J\times M,M)$, $h_0=\mathrm{id}$, $h_1$ 是 $T(G)$-等变的, 则 $i(A)\leqslant i(\overline{h_1A})$; (iv) 连续性: $A$ 有邻域 $N\in\sigma_G$, 使 $i(A)=i(N)$; (v) $i(A)=0\Leftrightarrow A=\varnothing$; 若 $0<|A|<\infty$, $|A|$ 记 $A$ 之基数, 则当 $|G|=1$, 或 $|G|>1$ 而 $0\overline{\in}A$ 时 $i(A)=1$. 以上 $A,B\in\sigma_G$. 当 $|G|=1$ 时就称 $i$ 为指标.

以下设 $i$ 是给定于 $M$ 上的一 $T(G)$-指标. 由指标的抽象定义可推出某些简单结论.

**7.6.2 命题** 设 $A\in\sigma_G$, 当 $|G|>1$ 时 $0\overline{\in}A$. 则 $i(A)\leqslant|A|$; 若 $A$ 紧且 $|G|<\infty$, 则 $i(A)<\infty$.

**证** 显然 $i(A)\leqslant|A|$. 若 $A$ 紧, $|G|<\infty$, 则 $\forall\ x\in A:i(G\cdot x)=1$. 取 $G\cdot x$ 的邻域 $N_x\in\sigma_G:i(N_x)=1$; 然后用一有限覆盖得出 $i(A)<\infty$. □

以下设 $f\in C^1(M)$ 满足(PS). 令 $\sigma_n=\{A\in\sigma_G:i(A)\geqslant n\}$, $c_n=c_{\sigma_n}$ 依 §4(1), 即

$$c_n=\inf\{\sup f(A):A\in\sigma_G,i(A)\geqslant n\},n\geqslant 1,\tag{1}$$

则 $\{c_n\}$ 就是本节开头所提到的增序列. 对 $c_n$ 所能直接验证的性质汇

集于下.

**7.6.3 命题**　设 $f$ 为 $T(G)$-不变，则(i) $-\infty\leqslant c_1\leqslant c_2\leqslant\cdots\leqslant c_n\leqslant\infty$；(ii) $c_1=\inf f(M)<\infty$，$c_1>-\infty\Leftrightarrow f$ 下有界；(iii) $i(M)<n\Rightarrow c_n=\infty$；(iv) $i(M)\geqslant n\Rightarrow c_n\leqslant\sup f(M)$，$i(M)=\infty\Rightarrow\forall\, n\geqslant 1$：$c_n\geqslant\sup f(M)$；(v) $c<c_n\Rightarrow i(f_c)<n$.

为建立本节的主要结果，需要 7.3.6 的以下推广：

**7.6.4 定理**　设 $f$ 为 $T(G)$-不变，$N\in\sigma_G$ 是 $K_c$ 的邻域，则存在 $\rho>r>0$，$h\in C(J\times M,M)$，使得 $h_0=\mathrm{id}$，$\forall\, t\in J$：$h_t\in\mathrm{Homeo}(M)$ 是 $T(G)$-等变的，$h_1(f_{c+r}\backslash N^\circ)\subset f_{c-r}$，当 $x\in M\backslash f^{-1}(c-\rho,c+\rho)$ 时 $h(t,x)=x$.

**证**　设 $\mu$ 是 $G$ 上的 Haar 测度(参考[75;Ch. 8])，$\mu G=1$；$\xi$ 是 $f$ 的伪梯度场. 约定 $a\cdot\zeta=dT_a(\zeta)$ $(a\in G,\zeta\in TM)$. 定义

$$\eta(x)=\int_G s^{-1}\cdot\xi(s\cdot x)\mathrm{d}s,x\in\tilde{M},\tag{2}$$

其中 $\mathrm{d}s$ 记 $\mathrm{d}\mu(s)$. $\eta$ 是 $\tilde{M}$ 上的向量场；$\forall\, a\in G$，由 $\mu$ 的不变性及链规则(§1(2))得出：

$$\begin{aligned}\eta(a\cdot x)&=\int_G s^{-1}\cdot\xi(s\cdot a\cdot x)\mathrm{d}s\\&=\int_G a\cdot(sa)^{-1}\cdot\xi(sa\cdot x)\mathrm{d}s=a\cdot\eta(x).\end{aligned}\tag{3}$$

因 $\xi\in C^{1-0}$ 而 $G$ 紧，$\forall\, x\in\tilde{M}$，有 $G\cdot x$ 在 $M$ 上的邻域 $U$：$\mathrm{Lip}(\xi|U)<\infty$. 令 $V_\delta=M\cap B_\delta(x)$；取充分小的 $\delta>0$，使 $G\cdot V_\delta\subset U$，则对任给 $y,z\in V_\delta$ 有

$$\begin{aligned}|\eta(y)-\eta(z)|&\leqslant\int_G|s^{-1}\cdot\xi(s\cdot y)-s^{-1}\cdot\xi(s\cdot z)|\mathrm{d}s\\&=\int_G|\xi(s\cdot y)-\xi(s\cdot z)|\mathrm{d}s\leqslant\mathrm{const}|y-z|,\end{aligned}$$

可见 $\eta\in C^{1-0}$. 其次，结合(2)与§2(1)及 $df\circ T_a=df$ 有

$$\langle\mathrm{d}f(x),\eta(x)\rangle=\int_G\langle\mathrm{d}f(x),s^{-1}\cdot\xi(s\cdot x)\rangle\mathrm{d}s$$

$$= \int_G (\mathrm{d}f(s \cdot x), \xi(s \cdot x))\mathrm{d}s \geqslant \int_G |\mathrm{d}f(s \cdot x)|^2 \mathrm{d}s$$

$$= |\mathrm{d}f(x)|^2 = [\int_G |\mathrm{d}f(s \cdot x)|\mathrm{d}s]^2 \geqslant \frac{1}{4}|\eta(x)|^2,$$

这表明 $\eta$ 是 $f$ 的伪梯度场. $\eta$ 与 $\xi$ 不同之处在于 $\eta$ 满足(3)(这是关键所在!). 以 $\eta$ 替代 $\xi$,重复 7.3.6 的证明,可得定理所要求的 $h$. □

现在已可建立本节的中心结果. 为记号简便,约定$K_{nm} = K \cap f^{-1}[c_n, c_m](1 \leqslant n \leqslant m)$, $c_n$ 依(1). 基本的问题是给出基数 $|K_{nm}|$ 的一个估计.

**7.6.5 Lusternik-Schnirelman 定理** 设 $f \in C^1(M)$ 满足(PS)且 $T(G)$-不变,$-\infty < c_n \leqslant c_m < \infty$. (i)若 $m = n$,或 $c_n < c_m$, $|G| = 1$,或 $c_n < c_m$, $0 \overline{\in} K_{nm}$,则 $|K_{nm}| \geqslant m - n + 1$. (ii)若 $c_n < c_m$, $|G| > 1$, $0 \overline{\in} K_{nm}$, $\forall\, x \in K \backslash \{0\}$: $|G \cdot x| = |G|$,则 $|K_{nm}| \geqslant (m - n + 1)|G|$. (iii)若 $n < m$, $c_n = c_m$, $|G| = 1$ 或 $0 \overline{\in} K_{nm}$,则 $|K_{nm}| = \infty$. (iv)若 $|G| = 1$ 或 $0 \in K_{nm}$,则 $|K_{nm}| \geqslant m - n + 1$.

**证** 设 $n \leqslant m$, $c = c_n = c_m$. 取 $K_c$ 的邻域 $N \in \sigma_G$,使 $i(K_c) = i(N)$. 设 $h, r$ 如 7.6.4. 取 $A \in \sigma_G$: $i(A) \geqslant m$, $f(A) < c + r$. 利用 7.6.4, 7.6.3(v)及 $i$ 的性质有

$$\begin{aligned} m &\leqslant i(A) \leqslant i(A \backslash N^\circ) + i(N) \\ &\leqslant i(h_1(A \backslash N^\circ)) + i(N) \\ &\leqslant i(h_1(f_{c+r} \backslash N^\circ)) + i(N) \\ &\leqslant i(f_{c-r}) + i(N) < n + i(K_c), \end{aligned}$$

可见 $i(K_c) > m - n$,从而 $K_c \neq \varnothing$. 当 $n < m$, $|G| = 1$ 或 $0 \overline{\in} K_c$ 时 $|K_c| = \infty$,这得出(iii). 其次设 $c_n < c_m$. 若 $c_n < c_{n+1} < \cdots < c_m$,则 $|K_{nm}| = \sum_n^m |K_{c_i}| \geqslant m - n + 1$. 若对某个 $k$: $n \leqslant k < m$, $c_k = c_{k+1}$,则当 $|G| = 1$ 或 $0 \overline{\in} K_{nm}$ 时 $|K_{c_k}| = \infty$,从而 $|K_{nm}| = \infty$. 这得出(i). 余下结论是明显的. □

**7.6.6 推论** 设 $f, i, c_n$ 如 7.6.5,假定 $f$ 下有界. 则(i)$c_1$ 是 $f$ 的

临界值.(ii)若 $f(K)$ 有界,且有最大的 $n$ 使 $c_n<\infty$,则 $i(M)=n$.(iii)若 $|G|=1$,或 $0\overline{\in}K$,则 $|K|\geqslant i(M)$.

**证** (i)只需注意 $c_1=\inf f(M)$ 有限(7.6.3).

(ii)显然 $i(M)\geqslant n$. 取 $a\in\mathrm{R}$;$K\subset f_{a-1}$,则 $M$ 可强形变收缩到 $f_a$(7.3.2),于是 $i(M)=i(f_a)$(7.6.1(iii)),因此 $i(M)=n$(否则 $i(f_a)\geqslant n+1, a\geqslant\sup f(f_a)\geqslant c_{n+1}=\infty$!).

(iii)可设 $|K|<\infty$. 因 $c_n<\infty\Rightarrow n\leqslant|K|$(7.6.5),故有最大的 $n$ 使 $c_n<\infty$. 由(ii)有 $i(M)=n$,因此 $|K|\geqslant n=i(M)$. □

7.6.5 及其推论看来是深刻的;考虑到它仅用到很一般的假定,它的结论更异乎寻常. 然而我们立刻发现它几乎不能直接应用. 显然,有效的应用依赖于指标 $i$ 及 $c_n$ 的计算或估计,而这正是困难之所在,且只有对具体的指标才可能加以考虑. 下节考虑拓扑指标的两个实例,用以获得 7.6.5 的两个典型的具体化.

## §7 畴数与亏格

**7.7.1 定义** 设 $A\subset M$. 若存在同伦 $h\in C(J\times A,M)$,使 $h_0=\mathrm{id}, h_1(A)=x_0\in M$,则说 $A$ 在 $M$ 中可缩,并说 $h$ 将 $A$ 缩为 $x_0$. 若 $A$ 闭,则称

$$\mathrm{cat}A\triangleq\min\{n: A \text{ 可用 } M \text{ 中 } n(0<n\leqslant\infty) \text{ 个可缩闭集覆盖}\} \tag{1}$$

为 $A$ 在 $M$ 中的畴数,亦记作 $\mathrm{cat}_M A$;约定 $\mathrm{cat}\varnothing=0$.

**注** 1°利用 $M$ 为 ANR 可以指明(参考[43; §27]),闭集 $A\subset M$ 可缩 $\Leftrightarrow\exists\, h\in C(J\times M,M): h_0=\mathrm{id}, h_1(A)=x_0\in M$.

2° 从拓扑的观点看来,可缩集是简单或平凡的;因此可以说,在一定意义上畴数 $\mathrm{cat}A$ 量度了 $A$ 的"拓扑复杂性".

3° $\mathrm{cat}_M A$ 强烈地依赖于 $A$ 的"外围空间"$M$. 显然 $M\subset N\Rightarrow\mathrm{cat}_M A\geqslant\mathrm{cat}_N A$. 若 $M$ 本身可缩(如 $M=X$ 时即如此),则对任何闭集 $A\subset M$

有 $\mathrm{cat}_M A \leqslant \mathrm{cat}_M M = 1$. 在这种平凡的情况下，畴数概念没有实质意义.

**7.7.2 定理**(Palais[133],1966) $\mathrm{cat}_M$ 是 $M$ 上的指标；它在如下意义上是最大的：若 $i$ 是 $M$ 上的指标，则对任给闭集 $A \subset M$ 有 $i(A) \leqslant \mathrm{cat}_M A$.

**证** 首先验证 $\mathrm{cat} = \mathrm{cat}_M$ 有 7.6.1 中的性质(i)～(v). (i)(ii)是明显的. 设 $A \subset M$ 闭，$h \in C(J \times M, M)$，$h_0 = \mathrm{id}$，$\mathrm{cat}\, \overline{h_1 A} = n$，今证 $\mathrm{cat} A \leqslant n$(这得出(iii)). 可设 $n < \infty$. 取闭集 $A_j \subset M$，同伦 $h^j \in C(J \times M, M)$，使 $h_1(A) \subset \bigcup A_j$，$h_0^j = \mathrm{id}$，$h_1^j(A_j) = x_j \in M (1 \leqslant j \leqslant n)$. 令 $k_t^j = h_t^j \circ h_t (t \in J)$，则 $k_0^j = \mathrm{id}$，$k_1^j h_1^{-1}(A_j) = x_j$，$A \subset \bigcup h_1^{-1}(A_j)$，这表明 $\mathrm{cat} A \leqslant n$. 为证(iv)，可设 $\mathrm{cat} A < \infty$，进而可设 $\mathrm{cat} A = 1$(否则考虑 $A$ 的一个覆盖)，于是有 $h \in C(J \times A, M)$：$h_0 = \mathrm{id}$，$h_1(A) = x_0 \in M$. $h$ 显然可连续地延拓到集 $S \triangleq J \times A \cup \{0,1\} \times M$ 上，使 $h(o,x) \equiv x$，$h(1,x) \equiv x_0$. 利用 $J \times M$ 为 ANR，进而可将 $h$ 连续地延拓到 $S$ 的一邻域 $V$ 上. 由 $J \times A \subset V$ 与 $J$ 紧推出：存在 $A$ 的邻域 $N$，使 $J \times N \subset V$，显然 $\mathrm{cat} \overline{N} = 1$. 对于(v)，设 $A = \{x_1, \cdots, x_n\} \subset M$，则有 $M$ 中连结 $x_1 \cdots, x_n$ 的路 $L$(回忆本章已假定 $M$ 连通!)，$L$ 显然可缩，因此 $\mathrm{cat} A = 1$.

其次设 $i$ 是 $M$ 上任一指标，$A \subset M$，$\mathrm{cat} A = n$，今证 $i(A) \leqslant n$. 不妨设 $0 < n < \infty$，于是 $A \subset \bigcup A_j$，$A_j \subset M (i \leqslant j \leqslant n)$ 是可缩闭集. 设 $h^j$ 将 $A_j$ 缩为 $x_j \in M$，则

$$i(A) \leqslant \sum_j i(A_j) \leqslant \sum_j i(\{x_j\}) = n.$$

这就证得 cat 的极大性. □

**注** 由 7.6.6 可见，$i(M)$ 越大，使 $c_n < \infty$ 的 $c_n$ 就越多，因而由 $\{c_n\}$ 提供的临界点信息亦愈多. 这一点更明显地体现于 7.6.6 中的不等式 $|K| \geqslant i(M)$. 因此可以说，畴数作为指标是“最优的”. 下面考虑的另一指标“亏格”亦具这种“最优性”(看 7.7.4).

以下设 $M$ 是 $X$ 的 $C^r$ 子流形，假定 $M$ 是对称的，即 $M = -M$. 令

$\sigma_M=\{A\subset M: A=\overline{A}=-A\}$.

**7.7.3 定义** 任给 $A\in\sigma_X$，称

$$\mathrm{gen}A \triangleq \min\{n: 存在奇映射\ \varphi\in C(A,\mathrm{R}^n\backslash\{0\})\} \tag{2}$$

为 $A$ 的亏格(genus)，当(2)中的 $n$ 不存在时令 $\mathrm{gen}A=\infty$；约定 gen $\varnothing=0$.

**注** 1°若 $\varphi\in C(A,\mathrm{R}^n)$是奇映射而 $0\in A$，则必 $\varphi(0)=0$. 因此对 $0\in A\in\sigma_X$ 恒有 $\mathrm{gen}A=\infty$.

2°若 $A\in\sigma_X,\varphi\in C(A,\mathrm{R}^n\backslash\{0\})$是奇映射，则 $\varphi$ 有扩张 $\tilde{\varphi}\in C(X,\mathrm{R}^n)$(1.2.5)，于是 $\varphi_1\triangleq\mathrm{\hat{o}dd}\tilde{\varphi}$ 是奇映射且亦为 $\varphi$ 的扩张. 其次注意 $|\varphi(x)|^{-1}\varphi(x)\in C(A,S^{n-1})$. 因此，对于 $\varnothing\neq A\in\sigma_X$ 有：$\mathrm{gen}A\leqslant n\Leftrightarrow$存在奇映射 $\varphi\in C(X,\mathrm{R}^n)$使 $0\overline{\in}\varphi A\Leftrightarrow$存在奇映射 $\varphi\in C(A,S^{n-1})$.

亏格概念最早由 Yang 于 1955 年引入，后经 Krasnoselskii 与 Coffman 等人的改进逐步演化成现今所呈现的面貌.

**7.7.4 定理** gen 是一 $T(\mathrm{Z}_2)$-指标，且有性质

(P) $A\in\sigma_X,\varphi\in C(A,X)$为奇映射$\Rightarrow\mathrm{gen}A\leqslant\mathrm{gen}\,\overline{\varphi A}$.

gen 是有性质(P)的最大 $T(\mathrm{Z}_2)$-指标(与 7.7.2 对照).

**证** 首先验证 gen 有 7.6.1 中的性质(i)～(v). (i)是明显的. 以下设 $A,B\in\sigma_X,\varphi\in C(X,\mathrm{R}^n),\psi\in C(X,\mathrm{R}^m)$，$\varphi$ 与 $\psi$ 为奇映射，$O\overline{\in}\varphi A\cup\psi B$. 因 $h\triangleq(\varphi,\psi)\in C(X,\mathrm{R}^{m+n})$是奇映射且 $O\overline{\in}h(A\cup B)$，故 gen $(A\cup B)\leqslant m+n$，这得出(ii). 直接看出 gen 有性质(P)，而(P)蕴涵了(iii). 设 $\varphi A\subset S^{n-1}$，则 $N\triangleq\{x:|\varphi(x)|\geqslant1/2\}$是 $A$ 的邻域，显然 $N\in\sigma_X$，$\mathrm{gen}N\leqslant n$，这得出(iv). 若 $0\overline{\in}A=\{\pm x_j:1\leqslant j\leqslant n\}$，则由 $\varphi(\pm x_j)=\pm1$ 定义出奇映射 $\varphi\in C(A,S^0)$，因此 $\mathrm{gen}A=1$，这得出(v).

其次设 $i$ 是有性质(P)的任一 $T(\mathrm{Z}_2)$-指标，$A\in\sigma_X$，$\mathrm{gen}A=n$，今证 $i(A)\leqslant n$. 不妨设 $0<n<\infty$，于是有奇映射 $\varphi\in C(A,S^{n-1})$，以$\{e_j\}$记 $\mathrm{R}^n$ 的标准基，令 $P_j(\sum\lambda_k e_k)=\lambda_j$，$A_j=\{x\in A:|P_j\varphi(x)|\geqslant1/\sqrt{n}\}$，则 $A=\bigcup A_j$，$A_j\in\sigma_X$，$P_j$ 作为投影是奇映射. 因此

$$\varphi_j : A_j \to \{\pm e_j\}, x \mapsto (\mathrm{sgn} P_j \varphi(x)) e_j$$

是连续奇映射，这推出 $i(A) \leqslant \sum_j i(A_j) \leqslant \sum_j i(\{\pm e_j\}) = n$. □

对于足够"简单"的集 $A$，可利用定义式(1)(2)求出 $\mathrm{cat}A$ 与 $\mathrm{gen}A$. 下举一例.

**7.7.5 命题** 设 $B = B_1(0) \subset X, S = \partial B$，则 $\mathrm{gen}S = \dim X$；当 $\dim X < \infty$ 时 $\mathrm{cat}_S S = 2$，否则 $\mathrm{cat}_S S = 1$.

**证** 首先设 $X = \mathrm{R}^n$. 显然 $\mathrm{gen}S \leqslant n, \mathrm{cat}_S S \leqslant 2$(球面 $S$ 的"南"、"北"两半都在 $S$ 上可缩). 若 $\mathrm{gen}S < n$，则有奇映射 $\varphi \in C(S, S^{k-1}), k < n$，这与 1.8.6 矛盾. 因此 $\mathrm{gen}S = n$. 若 $\mathrm{cat}_S S = 1$，则存在 $h \in C(J \times S, S) : h_0 = \mathrm{id}, h_1(S) = x_0 \in S$，不妨设 $h$ 已扩张为 $h \in C(J \times \overline{B}, \mathrm{R}^n)$ (1.2.5). 任取 $e \in S \setminus \{x_0\}$，有 $\mathrm{R}_+ e \cap h_1(S) = \varnothing$，于是 $d(h_1, B, 0) = 0$ (1.3.5)，但这与 $d(h_0, B, 0) = 1$ 矛盾. 因此 $\mathrm{cat}_S S = 2$.

其次设 $\dim X = \infty$. $\forall\, n \geqslant 1$，取 $n$ 维子空间 $X_n \subset X$，则 $\mathrm{gen}S \geqslant \mathrm{gen}(X_n \cap S) = n$，这得出 $\mathrm{gen}S = \infty$. 取 $P \in C(\overline{B}, S)$，使 $P | S = \mathrm{id}$(参考[43;8.7])；令 $h(t,x) = P((1-t)x), (t,x) \in J \times S$，则 $h$ 将 $S$ 缩为 $P(0) \in S$，因此 $\mathrm{cat}_S S = 1$. □

以上结论显示出 gen 与 cat 差异颇大. 尽管不能从 7.7.5 推断一般地有 $\mathrm{gen}A \geqslant \mathrm{cat}A$，但在常见的情况下，利用 gen 能对集合作出更精细的分类. gen 与 cat 的另一重大差别是，对于 $A \in \sigma_M$，$\mathrm{gen}A$ 不依赖于 $M$，这使得使用 gen 较之 cat 更为方便. 直接看出，$\mathrm{gen} | \sigma_M$ 是 $M$ 上有性质(P)的最大 $T(\mathrm{Z}_2)$-指标.

以下设 $f \in C^1(M)$ 满足(PS)，令

$$c_n = \inf\{\sup f(A) : A = \overline{A} \subset M \text{ 且 } \mathrm{cat}_M A \geqslant n\}; \tag{3}$$

$$c'_n = \inf\{\sup f(A) : A \in \sigma_M \text{ 且 } \mathrm{gen}A \geqslant n\}. \tag{4}$$

相应地令 $K_{nm} = K \cap f^{-1}[c_n, c_m], K'_{nm} = K \cap f^{-1}[c'_n, c'_m], 1 \leqslant n \leqslant m$. 结合 7.6.5, 7.6.6, 7.7.2 与 7.7.4 得到两个具体化的 Lusternik-Schnirelman 定理：

**7.7.6 定理** 若 $n\leqslant m$，$c_n$ 与 $c_m$ 有限，则

$$|K_{nm}|\geqslant\begin{cases}m-n+1, & m=n \text{ 或 } c_n<c_m;\\ \infty, & n<m \text{ 且 } c_n=c_m.\end{cases} \tag{5}$$

若 $f$ 下有界，则 $|K|\geqslant\mathrm{cat}M$.

**7.7.7 定理** 设 $M$ 是 $X$ 的 $C^r$ 子流形，$M\in\sigma_X$，$f$ 是偶泛函，$0\overline{\in} K$，$n\leqslant m$，$c'_n$ 与 $c'_m$ 有限，则

$$|K'_{nm}|\geqslant\begin{cases}m-n+1, & m=n\\ 2(m-n+1), & c'_n<c'_m;\\ \infty, & n<m \text{ 且 } c'_n=c'_m.\end{cases} \tag{6}$$

若 $f$ 下有界，则 $|K|\geqslant\mathrm{gen}M$.

**7.7.8 例** 设 $X$ 是 Hilbert 空间，于是单位球面 $S$ 是 $X$ 的 $C^\infty$ 子流形；设 $f\in C^1(S)$ 满足(PS)且下有界. 若 $\dim X=\infty$，则结合 7.7.5～7.7.7 得出：$f$ 至少有一个临界点，当 $f$ 为偶泛函时必有无限多个临界点. 其次设 $\dim X<\infty$，此时 $S$ 紧，$f$ 自动地满足(PS)且下有界. 由 $\mathrm{cat}_S S=2$ 推出 $-\infty<c_1\leqslant c_2<c_3=\infty$，由 7.7.6 推出：$c_1,c_2$ 是 $f$ 的临界值；当 $c_1=c_2$ 时 $f$ 有无限多个临界点；当 $c_1<c_2$ 时 $f$ 可能仅有两个临界点(此由函数 $f(x_1,x_2,x_3)=x_3((x_1,x_2,x_3)\in S^2)$ 可见). 若 $f$ 为偶泛函，则 $-\infty<c'_1\leqslant c'_n<\infty$，由 7.7.7 推出：$c'_1,\cdots,c'_n$ 是 $f$ 的临界值；当某个 $c'_i=c'_{i+1}(1\leqslant i<n)$ 时 $f$ 有无限多个临界点；当 $c'_1<c'_n$ 时 $f$ 至少有 $2n$ 个临界点，即 $|K|\geqslant 2n$. 注意函数

$$f:S^1\to\mathbf{R},(\cos t,\sin t)\mapsto\sin^2 t$$

恰有 4 个临界点，可见不等式 $|K|\geqslant 2n$ 不能再改进.

以上例子表明，利用 7.7.7 比 7.7.6 能获得更多的临界点. 这首先是因为 7.7.7 中 $f$ 为偶泛函这一重要因素起了作用. 同时，如前面提到的，也因为与对称性相结合的指标 gen 包含了更多的信息.

## §8 非线性特征值问题

本节设 $X$ 为无限维可分自反空间. 给定 $f,g\in C^1(X)$，考虑方程

$$f'(x) = \lambda g'(x). \tag{1}$$

若 $X$ 是 Hilbert 空间，$A=A^*\in L(X)$，$f(x)=2^{-1}(Ax,x)$，$g(x)=2^{-1}|x|^2$，则(1)是一个线性特征值方程：

$$Ax = \lambda x. \tag{2}$$

若 $\alpha$ 是 $g$ 的正则值且 $M_\alpha \triangleq g^{-1}(\alpha)\neq\varnothing$，则 $M_\alpha$ 是 $X$ 的 $C^1$ 子流形，$\forall\ x\in M_\alpha$：$T_xM_\alpha=N(g'(x))$(7.1.4)；$\varphi_\alpha\triangleq f|M_\alpha\in C^1$(7.1.5)。固定一个如上的 $\alpha$，且就写 $M_\alpha,\varphi_\alpha$ 为 $M,\varphi$. 由 7.1.5，$\forall\ x\in M$，有 $\lambda\in\mathbb{R}$ 使$(\lambda,x)$满足(1)$\Leftrightarrow x$ 是 $\varphi$ 的临界点. 这就将解方程(1)的问题归于一个临界点问题，因而可考虑应用前几节的结果. 其次注意到，当$(\lambda,x)$满足(1)时，$\lambda$ 通常已由 $x$ 确定. 例如，若$\langle g'(x),x\rangle\neq 0$，则从(1)得出

$$\lambda = \lambda(x) = \langle f'(x),x\rangle/\langle g'(x),x\rangle. \tag{3}$$

若 $X$ 是 Hilbert 空间，$x\in M$，则从(1)得出

$$\lambda = (f'(x),g'(x))/|g'(x)|^2. \tag{4}$$

若进而假定 $g(x)=2^{-1}|x|^2$，$|x|=1$，则(3)(4)两式都得出 $\lambda=(f'(x),x)$；若再设 $f(x)=2^{-1}(Ax,x)$，$A=A^*\in L(X)$，则有 $\lambda=(Ax,x)=2f(x)$.

以下固定 $M=g^{-1}(\alpha)$，$\alpha$ 是 $g$ 的正则值；记 $\varphi=f|M$. 任给 $x\in M$，若$\langle g'(x),x\rangle\neq 0$，则 $x\overline{\in} N(g'(x))=T_xM$，于是有分解 $X=T_xM\oplus\mathbb{R}x$. 约定 $P_x:X\to T_xM$ 记投影，且令 $z_x=P_xz(z\in X)$. 容易直接验知

$$z_x = z - \langle g'(x),x\rangle^{-1}\langle g(x),z\rangle x; \tag{5}$$

$$|P_x| \leqslant 1 + |\langle g'(x),x\rangle|^{-1}|g'(x)||x|. \tag{6}$$

**7.8.1 引理** 设 $x\in M$，$\langle g'(x),x\rangle\neq 0$，$\lambda(x)$依(3)，令

$$Fx = f'(x) - \lambda(x)g'(x), \tag{7}$$

则 $Fx=\mathrm{d}\varphi(x)^\circ P_x$，且

$$|\mathrm{d}\varphi(x)| \leqslant |Fx| \leqslant |d\varphi(x)||P_x|. \tag{8}$$

特别，$\mathrm{d}\varphi(x)=0\Leftrightarrow Fx=0$.

**证** 任给 $z\in X$，用(3)(5)(7)直接算得

$$\begin{aligned}\langle Fx-\mathrm{d}\varphi(x)\circ P_x,z\rangle &= \langle Fx,z\rangle-\langle \mathrm{d}\varphi(x),z_x\rangle \\ &= \langle Fx-f'(x),z_x\rangle+\langle Fx,z-z_x\rangle=\langle Fx,z-z_x\rangle \\ &= \langle g'(x),x\rangle^{-1}\langle g'(x),z\rangle\langle f'(x)-\lambda(x)g'(x),x\rangle=0,\end{aligned}$$

可见 $Fx=\mathrm{d}\varphi(x)\circ P_x$. 显然 $|Fx|\leqslant|\mathrm{d}\varphi(x)||P_x|$. 其次，

$$|\mathrm{d}\varphi(x)|=\sup\{\langle Fx,z\rangle:z\in T_xM,|z|=1\}\leqslant|Fx|. \qquad \square$$

称 $T:X\to Y$ 强连续，若 $x_n\rightharpoonup x\Rightarrow Tx_n\to Tx$.

**7.8.2 引理** 若 $f'$ 强连续，则 $f$ 亦强连续，且 $f$ 与 $f'$ 皆为有界映射.

**证** 设 $x_n\rightharpoonup x$. 由中值定理有 $y_n\in[x_n,x]$，使

$$f(x_n)-f(x)=\langle f'(y_n),x_n-x\rangle. \tag{9}$$

因必 $y_n\rightharpoonup x$，故 $f'(y_n)\to f'(x)$，于是由(9)推出 $f(x_n)\to f(x)$，可见 $f$ 强连续. 任给有界集 $B\subset X$. 若 $f'(B)$ 无界，则 $\exists\, x_n\in B$: $|f'(x_n)|\geqslant n(n\geqslant1)$. 因 $X$ 自反，不妨设 $x_n\rightharpoonup x$，从而 $f'(x_n)\to f'(x)$，得出矛盾. 因此 $f'(B)$ 有界. 同理 $f(B)$ 亦必有界. $\square$

对 $\varphi$ 应用指标理论的关键在于验证 $\varphi$ 满足(PS). 然而这是一个很强的要求，通常只能部分地满足，以至难以直接应用 7.6.5 或 7.7.7. 为补偿"紧性"的不足，对 $c_n$（参考 §7(4)）的定义作如下修正：以 $\mathscr{C}$ 记 $M$ 的对称紧子集之全体，令

$$c_n\doteq\inf\{\sup f(A):A\in\mathscr{C}\text{ 且 }\operatorname{gen}A\geqslant n\}. \tag{10}$$

**7.8.3 定理** 设 $f,g\in C^1(X)$ 为偶泛函且满足条件：

($H_1$) $f(0)=0,x\neq0\Rightarrow f(x)<0$;

($H_2$) $f'$ 强连续，$x\neq0\Rightarrow f'(x)\neq0$;

($H_3$) $g(0)=0,|x|\to\infty\Rightarrow|g(x)|\to\infty$;

($H_4$) $g'$ 有界，$\alpha>0$，$\inf\limits_{g(x)=\alpha}\langle g'(x),x\rangle>0$;

($H_5$) $x_n\rightharpoonup x$ 且 $g'(x_n)\to u\Rightarrow x_n\to x$.

设 $M=g^{-1}(\alpha)\neq\varnothing$，$c_n$ 依(10). (i)若 $c_n<\infty$，则 $-\infty<c_n<0$，且$\exists\, x_n$

$\in M$，使$(\lambda_n, x_n)$满足(1)，此处$\lambda_n=\lambda(x_n)$，$\lambda(\cdot)$依(3). (ii)若补充条件

$(\mathrm{H}_6)$　$\sup\{\mathrm{gen}A : A\in\mathscr{C}\}=\infty$，

则(1)有无穷多解. (iii)若$\forall\ n\geqslant 1 : c_n<\infty$，则$c_n\to 0(n\to\infty)$.

**证**　条件$(\mathrm{H}_4)$推出$\alpha$是$g$的正则值，故$M$是$X$的$C^1$子流形且$\overline{M}=M=-M$. 由$(\mathrm{H}_3)$，$M$有界且$0\overline{\in}M$.

令$\varphi=f|M, a<0$，今验证$\varphi$在$\varphi_a$上满足(PS). 设

$$\{x_n\}\subset\varphi_a, f(x_n)\text{ 有界}, |\mathrm{d}\varphi(x_n)|\to 0, \tag{11}$$

今证$\{x_n\}$含收敛子列. 因$M$有界，不妨设$x_n\rightharpoonup x$. 由(6)与条件$(\mathrm{H}_4)$知$|P_{x_n}|$有界，于是由(8)与$|\mathrm{d}\varphi(x_n)|\to 0$推出$Fx_n\to 0$，从而由(7)有

$$f'(x_n)-\lambda(x_n)g'(x_n)\to 0(n\to\infty). \tag{12}$$

由条件$(\mathrm{H}_2)$有$f'(x_n)\to f'(x)$，于是由(12)有

$$\lambda(x_n)g'(x_n)\to f'(x). \tag{13}$$

由(3)与条件$(\mathrm{H}_4)$知$\lambda(x_n)$有界，故不妨设$\lambda(x_n)\to\lambda_0$. 因$f$亦强连续(7.8.2)，故$f(x_n)\to f(x)\leqslant a$，于是由$(\mathrm{H}_1)(\mathrm{H}_2)$有$x\neq 0, f'(x)\neq 0$；这结合(13)与$g'(x_n)$有界得出$\lambda_0\neq 0, g'(x_n)\to\lambda_0^{-1}f'(x)$，于是由条件$(\mathrm{H}_5)$得$x_n\to x$. 因此$\varphi$在$\varphi_a$上满足(PS).

由 7.8.2，$f(M)$有界，因此$c_n>-\infty$. 由$(\mathrm{H}_1)$及$0\overline{\in}M$有$f(M)<0$，故当$c_n$有限时$c_n<0$. 设$c=c_n<0$. 若$c$非$\varphi$之临界值，则对充分小的$\varepsilon>0$，$\varphi$在$\varphi_{c+\varepsilon}$上满足(PS)，因而仍可得出如 7.6.4 的$h$，$r$(对于$G=z_2$)，可设$0<r<\varepsilon$. 取$A\in\mathscr{C}$，使$\mathrm{gen}A\geqslant n, f(A)<c+r$. 因$h_1(A)\subset M$仍为紧对称集，且$\mathrm{gen}h_1(A)\geqslant\mathrm{gen}A\geqslant n$，故由(10)有

$$c\leqslant\sup f(h_1(A))\leqslant\sup f(\varphi_{c-r})\leqslant c-r,$$

得出矛盾. 因此$c$是$\varphi$之临界值，而这得出(i).

若条件$(\mathrm{H}_6)$满足，则$\forall\ n\geqslant 1$，$\exists\ A_n\in\mathscr{C}:\mathrm{gen}A_n\geqslant n$，从而$c_n\leqslant\sup f(A_n)<\infty$. 若$c_n(n\geqslant 1)$互不相同，则由(i)直接得出(1)有无穷多解. 若有$n<m : c_n=c_m$，则用类似于 7.6.5 的证法可得出$\varphi^{-1}(c_n)$含$\varphi$的无限多个临界点，从而(1)亦有无穷多解. 这得出(ii).

设$\forall\ n\geqslant 1: c_n<0$,今证$c_n\to 0$. 若$c_n\not\to 0$,则有$\varepsilon>0,\forall\ n\geqslant 1: c_n<-\varepsilon$. 取$A_n\in\mathscr{C}: \mathrm{gen}A_n\geqslant n, f(A_n)<-\varepsilon$. 设$P_k: X\to X_k$如下面补证的7.8.4,则有$k\geqslant 1$:

$$\sup_{x\in M}|f(x)-f(P_kx)|<\varepsilon/2. \tag{14}$$

否则,$\forall\ k\geqslant 1,\exists\ x_k\in M: |f(x_k)-f(P_kx_k)|\geqslant\varepsilon/2$. 不妨设$x_k\rightharpoonup x$,从而$P_kx_k\rightharpoonup x$. 由$f$强连续得$f(x_k)-f(P_kx_k)\to 0$,得出矛盾. $\forall\ n\geqslant 1$,由(14)推出

$$\sup_{x\in A_n}|f(P_kx)|\geqslant\sup_{x\in A_n}|f(x)|-\frac{\varepsilon}{2}>\frac{\varepsilon}{2},$$

可见$0\overline{\in}P_kA_n$. 于是

$$n\leqslant\mathrm{gen}A_n\leqslant\mathrm{gen}(P_kA_n)\leqslant\dim X_k,$$

得出矛盾. 这证得(iii). □

**7.8.4 引理** 若$X$是可分的自反空间,则有连续奇映射序列$P_k: X\to X_k, k=1,2,\cdots, X_k$是$X$的有限维子空间,使得$x_k\rightharpoonup x\Rightarrow P_kx_k\rightharpoonup x, |P_kx|\leqslant|x|$.

**证** 只考虑$X$为Hilbert空间的情况(参考[202;44.32]). 设$\{e_k\}$是$X$的标准正交基,令$X_k=\mathrm{span}\{e_1,e_2,\cdots,e_k\}, P_k: X\to X_k$是正投影. 若$x_k\rightharpoonup x$,则$\{P_kx_k\}$的任何子列有弱收敛子列. 不妨设$P_kx_k\rightharpoonup y$,由

$$\begin{aligned}|x-y|^2&=\lim_k(x_k-P_kx_k, x-y)\\&=\lim_k(x_k,(I-P_k)(x-y))=0\end{aligned}$$

得$x=y$,这推出$P_kx_k\rightharpoonup x$. □

**注** 关于7.8.3的一个更一般的形式,参看[202;Th. 44. A].

取某些具体的$f,g$,可得出7.8.3的一些推论. 首先取$g(x)=2^{-1}|x|^2$,则$g'(x)=\mathscr{D}x$(3.2.5). 利用对偶映射$\mathscr{D}$的性质(参考§3.2)与7.8.3得出:

**7.8.5 推论** 设$X$可分且$X$与$X^*$皆一致凸,$f\in C^1(X)$是偶泛

函且满足$(H_1)(H_2)$(依 7.8.3,下同),$M=\partial B_r(0)\subset X$. 则有无穷多个$(\lambda_n,x_n)\in \mathrm{R}\times M$ 满足

$$f'(x)=\lambda\mathscr{D}x, \tag{15}$$

其中 $\lambda_n=r^{-2}\langle f'(x_n),x_n\rangle$.

**证** 令 $g(x)=2^{-1}|x|^2$,则 $g'(x)=\mathscr{D}x,\mathscr{D}:X\to X^*$ 是同胚(3.2.5). 直接看出$(H_3)(H_4)$满足. 若 $x_n\rightharpoonup x,\mathscr{D}x_n\to u$,则 $x_n\to\mathscr{D}^{-1}u=x$,故$(H_5)$亦满足. 注意 $M=g^{-1}(r^2/2)$. 任给有限维子空间 $Y\subset X$, $A=M\cap Y$ 是 $M$ 的对称紧子集,$\mathrm{gen}A=\dim Y$,这推出条件$(H_6)$满足. 于是用 7.8.3(ii)得出所要结论. □

**7.8.6 推论** 设 $X$ 是无穷维可分 Hilbert 空间,$A=A^*\in CL(X)$,$x\neq 0\Rightarrow(Ax,x)<0$,

$$\lambda_n=2\inf\{\sup f(S\cap Y):Y\subset X,n\leqslant \dim Y<\infty\},$$

其中 $f(x)=2^{-1}(Ax,x)$,$S=\partial B_1(0)\subset X$. 则 $\lambda_n$ 是 $A$ 的负特征值,$\lambda_n=(Ax_n,x_n)$,$x_n\in S$;$\lambda_n\to 0$.

**证** 显然 $f$ 满足$(H_1)(H_2)$,$f'(x)=Ax$;而 $g(x)=2^{-1}|x|^2$ 满足$(H_3)\sim(H_5)$,$g'=I$. 令 $\lambda_n=2c_n$,则易验知 $c_n$ 与(10)一致. 于是所要结论由 7.8.3 得出. □

**注** 当$(Ax,x)>0(x\neq 0)$时以$-A$代 $A$ 得出类似结论.

## §9 奇异同调群

本章引言已经指出,在临界点理论中使用拓扑工具是不可避免的. 不过,迄今为止所涉及的拓扑概念还是很初等的. 进一步深入需要更多的拓扑知识,尤其是有关同调论的结果. 为便于引用,本节给出奇异同调论的一个梗概. 我们只是罗列了基本的定义与结论,证明与详细讨论请参看文献(如[126]).

本节中 $X,Y$ 记拓扑空间,$A\subset X,B\subset Y$,称$(X,A)$为一个"空间对". $G$ 记一个加群.

设$\{e_i\}$是$\mathbf{R}^n$的标准基. 称$\Delta_p = \text{co}\{0, e_1, \cdots, e_p\}$为"标准$p$-单形", 称任何$\sigma \in \Sigma_p \triangleq C(\Delta_p, X)$为$X$上的"奇异$P$-单形", 称任何形式线性组合$\Sigma\lambda_i\sigma_i(\lambda_i \in G, \sigma_i \in \Sigma_p)$为$p$-链, 其全体依自然的运算构成一加群$C_p(X,G)$. 令$C_p(X,A;G)=C_p(X,G)/C_p(A,G)$.

任给$\sigma \in \Sigma_p$, 定义其"边界"$\partial\sigma$为: $\partial\sigma = \sum_0^p(-1)^i\sigma_i, \sigma_i = \sigma | S_i \in \Sigma_{p-1}$, $S_i$是$\Delta_p$的第$i$个顶点所对的"面"(它是一个$(p-1)$-单形). 将$\partial$线性地扩张到$C_p(X,G)$上, 得到同态$\partial: C_p(X,G) \to C_{p-1}(X,G)$, 称之为边界算子, 亦记作$\partial_p$. 可验证$\partial^2=0$(这对于定义同调群是关键的!). $\partial$自然地导出一同态$\partial' : C_p(X,A;G) \to C_{p-1}(X,A;G)$, 且$\partial'^2 = 0$; $\partial'$亦记作$\partial'_p$. 称$H_p(X,G) \triangleq \text{Ker}\partial_p/\text{Im}\partial_{p+1}$为$p$阶奇异同调群, 称$H_p(X,A;G) \triangleq \text{Ker}\partial'_p/\text{Im}\partial'_{p+1}$为$p$阶奇异相对同调群, 约定$H_p(X, \varnothing; G) = H_p(X,G)$.

令$C^p(X,G) = \text{Hom}(C_p(X,G), G)$(从$C_p(X,G)$到$G$的同态群), $C^p(X,A;G)$仿此. 称$\partial$的对偶$\delta: C^p(X,G) \to C^{p+1}(X,G), c \mapsto c \circ \partial$为"上边界算子", 亦记作$\delta_p$; $\delta'$或$\delta'_p$的定义仿此. 由$\partial^2=0$与$\partial'^2=0$推出$\delta^2=0$与$\delta'^2=0$. 称$H^p(X,G) \triangleq \text{Ker}\delta_p/\text{Im}\delta_{p-1}$为$p$阶奇异上同调群, 称$H^p(X,A;G) \triangleq \text{Ker}\delta'_p/\text{Im}\delta'_{p-1}$为$p$阶奇异相对上同调群.

不致误解时在同调群记号中省去$G$. $H_p(X)$中的元素表成$[c]$, 即$[c] = c + \text{Im}\partial_{p+1}, c \in \text{Ker}\partial_p$. 对$H_p(X,A), H^p(X), H^p(X,A)$中的元用类似记号. 以$H_*(X)$记群列$\{H_p(X): p \geqslant 0\}$; $H_*(X,A), H^*(X), H^*(X,A)$仿此. 同态$h: H_*(X) \to H_*(Y)$意味着一列同态$h_p: H_p(X) \to H_p(Y)(p \geqslant 0)$; $H_*(X) \cong H_*(Y)$(同构)$\Leftrightarrow \forall\ p \geqslant 0: H_p(X) \cong H_p(Y)$. 以上概念自然推广到$H_*(X,A), H^*(X)$与$H^*(X,A)$.

约定$f \in C(X,A;Y,B) \Leftrightarrow f \in C(X,Y)$且$fA \subset B$.

**7.9.1 定理** 设$f \in C(X,A;Y,B), g \in C(Y,B;Z,C)$. 则$f$唯一地诱导出同态$f_*: H_*(X,A) \to H_*(Y,B)$与$f^*: H^*(Y,B) \to H^*$

$(X,A)$，使得 $\partial f_* = f_* \partial$，$\delta' f^* = f^* \delta'$；若 $f=\mathrm{id}$，则 $f_*=\mathrm{id}$，$f^*=\mathrm{id}$；$(g\circ f)_*=g_*\circ f_*$，$(g\circ f)^*=f^*\circ g^*$.

约定 $f\simeq g:(X,A)\to(Y,B)\Leftrightarrow\exists\, h\in C(J\times X,Y)$：$h_0=f$，$h_1=g$，$h(J\times A)\subset B$；$(X,A)\simeq(Y,B)\Leftrightarrow\exists\, f\in C(X,A;Y,B)$，$g\in C(Y,B;X,A)$，使得 $g\circ f\simeq\mathrm{id}:(X,A)\to(X,A)$，$f\circ g\simeq\mathrm{id}:(Y,B)\to(Y,B)$. 若 $h\in C(J\times X,X)$ 满足 $h_0=\mathrm{id}$，$h_1(X)\subset Y\subset X$，$h_1(A)\subset B$，$h_t(A)\subset A$，$h_t|Y=\mathrm{id}$，则说 $h$ 是从 $(X,A)$ 到 $(Y,B)$ 的一个形变收缩，此时必定 $(X,A)\simeq(Y,B)$ 且 $Y,B$ 分别为 $X,A$ 的强形变收缩核.

**7.9.2 定理**（同伦不变性） 若 $f\simeq g:(X,A)\to(Y,B)$，则 $f_*=g_*$，$f^*=g^*$；若 $(X,A)\simeq(Y,B)$（特别，若 $(X,A)$ 可形变收缩到 $(Y,B)$），则 $H_*(X,A)\cong H_*(Y,B)$，$H^*(X,A)\cong H^*(Y,B)$.

**7.9.3 切除定理** 若 $U\subset A^\circ$，则 $H_*(X\backslash U,A\backslash U)\cong H_*(X,A)$，$H^*(X\backslash U,A\backslash U)\cong H^*(X,A)$.

若一（有限或无限的）群同态序列

$$\cdots\to G_i\xrightarrow{h_i}G_{i+1}\xrightarrow{h_{i+1}}\cdots$$

满足 $\mathrm{Im}h_i=\mathrm{Ker}h_{i+1}$，则称此序列为恰当序列. 例如，"短序列" $0\to F\xrightarrow{\varphi}G\xrightarrow{\psi}H\to 0$ 是恰当的 $\Leftrightarrow\varphi$ 是单同态、$\psi$ 是满同态且 $\mathrm{Im}\varphi=\mathrm{Ker}\psi$；$0\to G\xrightarrow{\varphi}H\to 0$ 是恰当的 $\Leftrightarrow\varphi:G\cong H$，此处 0 记平凡群（下同）.

**7.9.4 定理** 若 $A\subset B\subset X$，则有恰当序列：

$$\cdots\to H_p(B,A)\xrightarrow{i_*}H_p(X,A)\xrightarrow{j_*}H_p(X,B)\xrightarrow{\partial}H_{p-1}(B,A)\to\cdots;\tag{1}$$

$$\cdots\to H_p(B)\xrightarrow{i_*}H_p(X)\xrightarrow{j_*}H_p(X,B)\xrightarrow{\partial}H_{p-1}(B)\to\cdots;\tag{2}$$

$$\cdots\leftarrow H^p(B,A)\xleftarrow{i^*}H^p(X,A)\xleftarrow{j^*}H^p(X,B)\xleftarrow{\delta}H^{p-1}(B,A)\leftarrow\cdots;\tag{3}$$

$$\cdots\leftarrow H^p(B)\xleftarrow{i^*}H^p(X)\xleftarrow{j^*}H^p(X,B)\xleftarrow{\delta}H^{p-1}(B)\leftarrow\cdots,\tag{4}$$

其中 $i:(B,A)\to(X,A)$ 与 $j:(X,A)\to(X,B)$ 是包含映射；(1)(2) 中的 $\partial$ 决定于 $\partial[c]=[\partial c]$；(3)(4) 中的 $\delta$ 仿此.

**7.9.5 定理** 设$\{X_k\}$是 $X$ 的路连通分支之全体，则 $H_p(X,A)\cong\bigoplus_k H_p(X_k,X_k\cap A)(p\geqslant 0)$.

应用同调论时必然涉及某些同调群的计算.直接依定义完成计算的希望很小，通常是利用少数已知结果与同调群的性质(如 7.9.1～7.9.5)间接地得出所求同调群.基本的已知结果如下：$H_p(X,X)=0$；若 $X$ 路连通，$A\neq\varnothing$，则 $H_0(X,A)=0$；若 $X$ 可强形变收缩为一点，则 $H_p(X)=0$，$H^p(X)=0(p\neq 0)$，$H^0(X,G)\cong G\cong H_0(X,G)$；

$$H_p(S^n,G)\cong\begin{cases}0, & p,n\geqslant 1,p\neq n;\\ G, & p=n\geqslant 1 \text{ 或 } p=0,n\geqslant 1;\end{cases}\tag{5}$$

$$H_p(B^n,B^n\setminus\{0\};G)\cong\begin{cases}0, & p\neq n;\\ G, & p=n,\end{cases}\tag{6}$$

其中 $B^n$ 是 $n$ 维单位球.

在同调群对临界点理论的应用中，通常取"系数群"$G$ 为某个域 $F$(如 $F=\mathrm{Z}_2$，R 或 C)，此时 $H_p(X,A,F)$实际上是 $F$ 上的向量空间.定义

$$\beta_p(X,A;F)=\dim H_p(X,A;F);\tag{7}$$

$$S_p(X,A;F)=\sum_{i=0}^{p}(-1)^{p-i}\beta_i(X,A;F);\tag{8}$$

$$\chi(X,A;F)=\sum_{p\geqslant 0}(-1)^p\beta_p(X,A;F)\quad,\tag{9}$$

(9)式要求右端级数收敛.称 $\beta_p$ 为 $p$ 阶 Betti 数，称 $\chi$ 为 Euler 特征数，它们是刻划空间对$(X,A)$拓扑特性的基本不变量.不致误解时，(7)～(9)式中字母 $F$ 可以省去.

**7.9.6 定理** 若 $A\subset B\subset X$，则

$$\beta_p(X,A)\leqslant\beta_p(X,B)+\beta_p(B,A);\tag{10}$$

$$S_p(X,A)\leqslant S_p(X,B)+S_p(B,A);\tag{11}$$

$$\chi(X,A)=\chi(X,B)+\chi(B,A)\quad,\tag{12}$$

(12)式要求其右端两项有定义.

## §10 Morse 理论

本节设 $X$ 是 Hilbert 空间,给定 $f\in C^2(X)$.

若 $x_0\in K, f''(x_0)\in GL(X)$,则称 $x_0$ 为 $f$ 的非退化临界点;当 $f$ 仅有非退化临界点时称 $f$ 为 Morse 函数. 用反函数定理容易推出,非退化临界点必为孤立临界点. 下面进而指明,$f$ 在其非退化临近点附近的性状如同二次函数.

**7.10.1 Morse 引理** 设 $x_0$ 是 $f$ 的非退化临界点,则有 $x=0$ 处的局部 $C^1$ 同胚 $g,h,g(0)=h(0)=x_0$,使得

$$f(g(x)) = f(x_0) + 2^{-1}(Ax,x); \tag{1}$$

$$f(h(x)) = f(x_0) + 2^{-1}(|x_+|^2 - |x_-|^2), \tag{2}$$

其中 $A=f''(x_0), x=x_++x_-, x_\pm\in X_\pm, X_+$ 与 $X_-$ 分别为 $A$ 关于正谱与负谱的不变子空间.

**证** 为简单起见,在 $f\in C^3$ 的假定下证明. 可设 $x_0=0, f(0)=0$. 令 $L_s(X)=\{T\in L(X):T=T^*\}$. 定义

$$B(x) = \int_0^1 (1-t)f''(tx)\mathrm{d}t, x\in X;$$

$$\Phi(T,x) = TA^{-1}T - 2B(x), (T,x)\in L_s(X)\times X.$$

直接看出 $\Phi\in C^1(L_s(X)\times X, L_s(X)), \Phi(A,0)=0, \Phi'_T(A,0)=2\mathrm{id}$. 由隐函数定理,方程 $\Phi(T,x)=0$ 在 $(A,0)$ 邻近有 $C^1$ 解 $T=T(x)$,即 $T(x)A^{-1}T(x)=2B(x), T(0)=A$. 令 $\varphi(x)=A^{-1}T(x)x$,则 $\varphi\in C^1$, $\varphi(0)=0, \varphi'(0)=I$. 由反函数定理,$\varphi$ 是 $x=0$ 处的局部 $C^1$ 同胚. 令 $g=\varphi^{-1}$,则 $g(0)=0, y=g(x)$ 满足:

$$f(g(x)) = f(y) = \int_0^1 (1-t)f''(ty)y^2\mathrm{d}t$$

$$= (B(y)y,y) = \frac{1}{2}(T(y)A^{-1}T(y)y,y)$$

$$= \frac{1}{2}(\varphi(y), A\varphi(y)) = \frac{1}{2}(Ax, x).$$

其次，令 $A_{\pm} = \pm A|X_{\pm}$，$S = \sqrt{A_+} \oplus (-\sqrt{A_-})$，$h = gS^{-1}$，则

$$f(h(x)) = \frac{1}{2}(AS^{-1}x, S^{-1}x) = \frac{1}{2}(|x_+|^2 - |x_-|^2). \qquad \square$$

本节的主要目的是用同调群来描述 $f$ 的临界点，所用的同调群都以某个域 $F$ 为系数群.

首先考虑 $f$ 在孤立临界点处的局部性态.

**7.10.2 定义** 设 $f$ 在开集 $U$ 内仅含临界点 $x_0$，$f(x_0)=c$. 称

$$C_p(x_0, f) \triangleq H_p(f_c \cap U, f_c \cap (U \setminus \{x_0\})) \tag{3}$$

为 $f$ 在 $x_0$ 的临界群. 若 $x_0$ 是非退化临界点，$X_-$ 如 7.10.1，则称 $\dim X_-$ 为 $f$ 在 $x_0$ 的 Morse 指数.

由切除定理(7.9.3)，(3)与 $U$ 的选取无关. 若 $x_0$ 是非退化的，取 $U$ 为一个球，则由(2)观察出 $(f_c \cap U, f_c \cap (U \setminus \{x_0\}))$ 有简单的形变收缩. 循此思路得到

**7.10.3 定理** 若 $x_0$ 是 $f$ 的非退化临界点，其 Morse 指数为 $j$，则 $C_p(x_0, f) = 0\ (p \neq j)$，$C_j(x_0, f) = F\ (j < \infty)$.

**证** 可设 $x_0 = 0$，$f(x) = 2^{-1}(|x_+|^2 - |x_-|^2)$（依 7.10.1 的记号）. 分别以 $B, B^j$ 记 $X, X_-$ 中的单位球，令 $h(t, x) = tx_- + (1-t)x$. 不难看出 $h$ 将 $(f_0 \cap B, f_0 \cap (B \setminus \{0\}))$ 形变收缩到 $(B^j, B^j \setminus \{0\})$，于是 $C_p(0, f) = H_p(B^j, B^j \setminus \{0\})$ (7.9.2)，当 $j < \infty$ 时由此立得定理结论. 若 $j = \infty$，不妨设 $B^j = B$，则有 $\varphi \in C(\overline{B}, \overline{B})$：$\mathrm{Fix}\varphi = \varnothing$ (1.4.4). $\forall x \in \overline{B}$，存在唯一 $\psi(x) \in S = \partial B$，使 $x \in [\varphi(x), \psi(x)]$；这样得到 $\psi \in C(\overline{B}, S)$，$\psi|S = \mathrm{id}$. 直接看出 $(t, x) \mapsto t\psi(x) + (1-t)x$ 是从 $(\overline{B}, \overline{B} \setminus \{0\})$ 到 $(S, S)$ 的形变收缩，因此

$$C_p(0, f) = H_p(B, B \setminus \{0\}) = H_p(S, S) = 0. \qquad \square$$

7.10.3 表明，非退化临界点的临界群完全决定于 Morse 指数. 若 $j = 0$，则由(2)知 $x_0$ 是 $f$ 的严格极小点，由 7.10.3 有 $C_0(x_0, f) =$

$F, C_p(x_0, f) = 0(p>0)$. 若 $X = \mathrm{R}^n, j = n$，则 $x_0$ 是 $f$ 的严格极大点，$C_n(x_0, f) = F, C_p(x_0, f) = 0(p \neq n)$.

为用同调群描述与临界点有关的整体性质，需用一定形变定理，此时不免用到(PS)条件.

**7.10.4 引理** 设 $f$ 满足(PS)，在$[a,b](a<b)$上仅有临界值 $c$，$|K_c|<\infty$，则 $H_*(f_b, f_a) \cong H_*(f_c, f_c \backslash K_c)$.

**证** 由 7.3.5 与 7.9.2 推出 $H_*(f_b, f_a) \cong H_*(f_c, f_a)$. 由 7.3.2 有 $H_*(f_c \backslash K_c, f_a) \cong H_*(f_a, f_a) = 0$，于是由 §9(1)有恰当序列

$$0 \to H_p(f_c, f_a) \to H_p(f_c, f_c \backslash K_c) \to 0,$$

由此得 $H_p(f_c, f_a) \cong H_p(f_c, f_c \backslash K_c)$，如所要证. □

若 $K_c = \{x_1, x_2, \cdots, x_n\}$，则用切除定理易推出

$$H_p(f_c, f_c \backslash K_c) = \bigoplus_i C_p(x_i, f). \tag{4}$$

设 $f$ 只有孤立临界值 $\cdots < c_{-1} < c_0 < c_1 < \cdots$，$|K_{c_i}| < \infty$，取 $\varepsilon_i > 0$ 使 $\varepsilon_i < \min\{c_{i+1} - c_i, c_i - c_{i-1}\}$，称

$$M_p \triangleq \sum_i \dim H_p(f_{c_i+\varepsilon_i}, f_{c_i-\varepsilon_i}) \tag{5}$$

为 $f$ 的 $p$ 阶 Morse 型数. 结合(3)与 7.10.3，7.10.4 得出：

**7.10.5 推论** 若 $f$ 是 Morse 函数且满足(PS)，则 $Mp =$"$f$ 的 Morse 指数为 $p$ 的临界点之个数".

**7.10.6 Morse 不等式** 设 $f$ 满足(PS)，$a<b, K_a = a = K_b, |K_{ab}| < \infty$，每个 $x \in K_{ab}$ 有有限维的临界群，$M_p$ 是 $f$ 在 $f^{-1}[a,b]$上的 Morse 型数，则

$$S_p(f_b, f_a) \leqslant \sum_{i=0}^{p} (-1)^{p-i} M_i; \tag{6}$$

$$\chi(f_b, f_a) = \sum_p (-1)^p M_p, \tag{7}$$

$S_p, \chi$ 依 §9(8)(9)，假定(7)右端收敛.

**证** 设 $f$ 在$(a,b)$内的临界值为$\{c_i\}$，$a = a_0 < c_1 < a_1 < c_2 < \cdots < c_n < a_n = b$，则由(5)与 §9(11)有

$$\sum_{i=0}^{p}(-1)^{p-i}M_i = \sum_{i=0}^{p}(-1)^{p-i}\sum_{j=1}^{n}\beta_i(f_{a_j},f_{a_{j-1}})$$

$$= \sum_{j=1}^{n}S_p(f_{a_j},f_{a_{j-1}}) \geqslant S_p(f_b,f_a),$$

即(6)成立. 其次,设 $\sum(-1)^p M_p$ 收敛,则∃ $p\geqslant 0,\forall\, i>p: M_i=0$. 令 $\beta_i=\beta_i(f_b,f_a)$,则结合(6)与 § 9(8)有

$$\beta_{p+1} = S_{p+1}(f_b,f_a) + S_p(f_b,f_a)$$

$$\leqslant \sum_{i=0}^{p}(-1)^{p+1-i}M_i + \sum_{i=0}^{p}(-1)^{p-i}M_i = 0,$$

由此得 $\beta_{p+1}=0$ 且 $S_p(f_b,f_a)=\sum_0^p(-1)^{p-i}M_i$. 同理,$\forall\, i>p: \beta_i=0$. 于是

$$\chi(f_b,f_a) = \sum_{i=0}^{p}(-1)^i\beta_i = (-1)^p\sum_{i=0}^{p}(-1)^{p-i}\beta_i$$

$$= (-1)^p S_p(f_b,f_a) = (-1)^p\sum_{i=0}^{p}(-1)^{p-i}M_i$$

$$= \sum_{i=0}^{\infty}(-1)^i M_i.$$ □

本节的局部性结果显然可推广于 Hilbert 流形上的 $C^2$ 函数(类比于 Banach 流形,Hilbert 流形的含义是自明的);而 7.10.4~7.10.6 基于局部性结果与形变定理,因此同样适用于 Hilbert 流形上的 $C^2$ 函数. 有鉴于此,不难从 7.10.6 得出:

**7.10.7 推论** 若 $M$ 是 $n$ 维 $C^2$ 紧流形,$f\in C^2(M)$ 为 Morse 函数,$M_i$ 是 $f$ 的 Morse 指数为 $i$ 的临界点之个数,$\beta_i=\dim H_i(M)$,则

$$\sum_{i=0}^{p}(-1)^{p-i}\beta_i \leqslant \sum_{i=0}^{p}(-1)^{p-i}M_i (0\leqslant p\leqslant n); \tag{8}$$

$$\sum_{i=0}^{n}(-1)^i\beta_i = \sum_{i=0}^{n}(-1)^i M_i. \tag{9}$$

**7.10.8 定理** 设 $f$ 满足(PS)且下有界,$x_0$ 是 $f$ 的非退化临界点,它非极小点且有 Morse 指数 $j<\infty$,则 $f$ 至少有 3 个临界点.

**证** 由 7.4.2，$f$ 有极小点 $\hat{x}$，可设 $\hat{x}$ 是唯一极小点．设 $f(\hat{x})=c<a<f(x_0)<b$，假定 $K=\{\hat{x},x_0\}$．由 § 9(12)有

$$\chi(f_b)=\chi(f_b,f_a)+\chi(f_a). \tag{10}$$

由(4)与 7.10.4 有 $H_p(f_b,f_a)=C_p(x_0,f)$；再结合 7.10.3 与 § 9(7)(9)得 $\chi(f_b,f_a)=(-1)^j$．类似地，由

$$H_p(f_a)=C_p(\hat{x},f)=H_p(\{\hat{x}\})=\begin{cases}F,p=0\\0,p>0\end{cases};$$

得 $\chi(f_a)=1$，因此 $\chi(f_b)=1+(-1)^j$．另一方面，由 7.3.2 知 $X$ 可强形变收缩到 $f_b$，这推出 $\chi(f_b)=\chi(X)=1$，得出矛盾．因此 $|K|\geqslant 3$. □

# 第八章　非线性动力系统

动力系统理论起源于19世纪末，关于常微系统的定性研究. 现代科学面对大量的时间动态系统，对这些系统中种种复杂现象进行定性考察的紧迫性，推动了动力系统理论的发展，并使其具有日益增长的重要性.

对于非线性动力系统的整体性质的研究，流形概念是必不可少的. 本章设 $M$ 是给定的 $C^rX$-流形，$r\geqslant 1$，$M$ 上已定义一 Finsler 结构 $|\cdot|$ 及由 $|\cdot|$ 导出的度量 $d$（参考§7.1）. 许多概念与结果并不依赖于 $\dim M$. 但一些细致的结果要求 $\dim M<\infty$.

## §1　基本用语

任给 $f\in \mathrm{Diff}^r(M)$，称变换群 $\{f^n:n\in \mathrm{Z}\}$ 为一个离散流. 任给 $M$ 上的 $C^r$ 向量场 $\xi$，当 $r\geqslant 1$ 时 $\xi$ 决定一（局部）流，约定记作 $\xi_t$，即 $\xi_t=\theta(t,\cdot)$，$t\mapsto\theta(t,x)$ 是 $\xi$ 过 $x$ 的积分曲线（7.2.3）. 如上的 $\{f^n\}$ 与 $\xi_t$ 分别称为离散动力系统与连续动力系统，概称动力系统. “离散”与“连续”这两种情况之间的类似比初看起来要深刻得多. 这一事实自从 Smale 在其著名的论文[168]中大加强调以来，已为众所周知. 动力系统理论的许多概念与结果，通常对形式上较简单的离散系统给定，然后过渡到连续系统；这种过渡虽然不总是平凡的，但通常有一些颇具启发性的思路可循. 因此，本章将以主要篇幅讨论离散系统.

以下给定 $f\in \mathrm{Diff}^r(M)$ 与 $M$ 上的 $C^r$ 向量场 $\xi$，$r\geqslant 1$. 下面一般只对 $f$ 给出有关术语与记号，通过代 $f^n$ 为 $\xi_t$，即得出关于 $\xi$ 的对应术语与记号.

**8.1.1 定义**　任给 $x\in M$，称 $\mathrm{O}(x)\triangleq\{f^n(x):n\in \mathrm{Z}\}$ 为 $f$ 过 $x$ 的

轨道；称 $O^+(x)\triangleq\{f^n(x):n\in Z_+\}$ 为正轨. 当 $n\to\infty$ 时 $\{f^n(x)\}$ 的任何极限点称为 $x$ 的 $\omega$ 极限点，其全体记作 $\omega(x)$. 令 $\mathrm{Per}f=\{x:\exists\ n\geqslant 1, f^n(x)=x\}$. 若 $x\in \mathrm{Per}f\setminus \mathrm{Fix}f$，则有最小的 $p>1:f^p(x)=x$，称这样的 $x$ 为 $p$-周期点，称 $O(x)$ 为 $p$-周期闭轨.

必要时以 $O(x,f)$ 记上述的 $O(x)$；$O^+(x,f)$，$\omega(x,f)$ 的含义仿此. 约定 $\alpha(x,f)=\omega(x,f^{-1})$，$L(f)=\bigcup\limits_{x\in M}(\omega(x,f)\cup\alpha(x,f))$. 显然 $\mathrm{Fix}f\subset \mathrm{Per}f\subset L(f)$，这三个集（$f$ 的不动点集、周期点集与极限点集）刻划了 $f$ 的不同程度的"回归性". 直接看出，$x\in \mathrm{Per}f\Leftrightarrow O(x)$ 紧 $\Leftrightarrow O(x)=\omega(x)$；若 $\overline{O^+(x)}$ 紧，则 $\omega(x)$ 是非空紧集.

对于 $\xi$ 完全类似地定义 $O(x)$，$O^+(x)$，$\omega(x)$，$\mathrm{Per}\xi$；$O(x)$ 亦写作 $O(x,\xi)$，余仿此. 与 $f$ 的不动点相当的是奇点：若 $\xi(x)=0$，则称 $x$ 为 $\xi$ 的奇点；实际上，$\xi(x)=0\Leftrightarrow x\in\bigcap\limits_t \mathrm{Fix}\xi_t$.

**8.1.2 命题** 关于 $\xi$ 有以下结论：(i) 若 $\overline{O^+(x)}$ 紧，则 $\omega(x)$ 是非空连通紧集；(ii) 若 $x\in \mathrm{Per}\xi$，$\xi(x)\neq 0$，则有最小的 $p>0:\xi_p(x)=x$（此时称 $O(x)$ 为 $p$-周期闭轨）；(iii) $x\in \mathrm{Per}\xi\Leftrightarrow O(x)$ 紧 $\Leftrightarrow O(x)=\omega(x)$.

**证** 只证(iii). 显然 $x\in \mathrm{Per}\xi\Rightarrow O(x)$ 紧 $\Rightarrow O(x)=\omega(x)$. 若 $O(x)=\omega(x)=\bigcup A_n$，$A_n=\{\xi_t(x):|t|\leqslant n\}$，则由 Baire 定理知某个 $A_n$ 在 $O(x)$ 中含内点 $y$. 设 $\xi_{t_i}(x)\to y(t_i\to\infty)$，则对大的 $i$ 有 $\xi_{t_i}(x)\in A_n$，这推出 $x\in \mathrm{Per}\xi$. □

动力系统理论的基本课题是阐明系统的轨道结构，这往往决定于对系统的某些"奇异"或"临界"元素——不动点与闭轨即属此类——的考察. 因此，不动点（或奇点）与闭轨有基本的重要性.

**8.1.3 定义** 设 $A\subset M$. 若 $\forall\ x\in A:O(x)\subset A[O^+(x)\subset A]$，则称 $A$ 为 $f$ 的不变集[正不变集].

显然，$A$ 对 $f$ 不变[正不变] $\Leftrightarrow fA=A[fA\subset A]$. $O(x)$，$\omega(x)$，$\mathrm{Fix}f$，$\mathrm{Per}f$，$L(f)$ 及下面定义的 $W^s(A)$ 等皆为不变集；$O^+(x)$ 是正不变集. 不变集的意义在于：若 $A$ 对 $f$ 不变，则可将"较小的"动力系统

$f|A$ 分离出来进行独立考察.

**8.1.4 定义** 设 $A$ 是 $f$ 的紧不变集. 分别称

$$W^s(A)=\{x:d(f^n(x),A)\to 0(n\to\infty)\} \tag{1}$$

与

$$W^u(A)=\{x:d(f^n(x),A)\to 0(n\to -\infty)\} \tag{2}$$

为 $A$ 的稳定集与不稳定集. 若 $V\subset W^s(A)$ 且 $\forall\ x\in V:\overline{O^+(x)}$ 紧, 则说 $A$ 吸引 $V$. 若 $W^s(A)$ 是 $A$ 的邻域, 则称 $A$ 为吸引子. 若 $x\in \mathrm{Fix}f$ 是吸引子, 对 $x$ 的任给邻域 $V$, 有 $x$ 的邻域 $U$, $\forall\ x\in U:O^+(x)\subset V$, 则说 $x$ 是渐近稳定不动点.

若能阐明 $W^s(A)$ 与 $W^u(A)$ 的构造, 则可认为对于系统在 $A$"邻近"的状态已得到描述. 在本章中, 对于 $A$ 仅含一"双曲不动点"与 $A$ 是"双曲闭轨"的情况给出了 $W^s(A)$ 与 $W^u(A)$ 的构造.

动力系统理论的基本观点本质上是拓扑的, 即它只关心系统的轨道结构, 不关心轨道的准确形状; 凡有相同轨道结构的系统不予区别. 问题在于如何界定轨道结构的异同, 这依赖于一套按"拓扑等价"分类的原则.

**8.1.5 定义** 设 $N$ 与 $M$ 一样为 $C^r$ 流形, $f\in \mathrm{Diff}^r(M)$, $g\in \mathrm{Diff}^r(N)$. 若 $A,B$ 分别为 $f$ 与 $g$ 的闭不变集, 存在同胚 $h:A\to B$, 使 $hf|A=gh$, 则说 $f$ 在 $A$ 上与 $g$ 拓扑共轭; 当 $A=M, B=N$ 时就说 $f$ 与 $g$ 拓扑共轭并记作 $f\sim g$. 若 $f:U\subset M\to M$ 与 $g:V\subset N\to N$ 分别在 $a\in U$ 与 $b\in V$ 处为局部 $C^r$ 同胚, 且有 $a$ 的邻域 $U_1\subset U$, $b$ 的邻域 $V_1\subset V$, 同胚 $h:U_1\cup fU_1\to V_1\cup gV_1$, 使 $h(a)=b$, $hU_1=V_1$, $hf|U_1=gh|U_1$, 则说 $f$ 与 $g$ 局部拓扑共轭, 记作 $f\overset{\mathrm{loc}}{\sim}g$, 当需指明 $a,b$ 时写作 $(f,a)\sim(g,b)$.

注意 $hf=gh\Rightarrow hf^n=g^nh$, 因此当 $f\sim g$ 时 $f$ 与 $g$ 的轨道在一同胚映射之下互相对应; 相应地, $f$ 与 $g$ 的不动点、闭轨等亦互相对应, 因而可以说有同样的轨道结构. 若 $f\overset{\mathrm{loc}}{\sim}g$, 则局部地有同样结论.

设 $\xi,\eta$ 分别为 $M$ 与 $N$ 上的 $C^r$ 向量场. 若有同胚 $h:M\to N$, 使得

$h\xi_t = \eta_t h$（这恰与 $hf^n = g^n h$ 对应！），则说 $\xi$ 与 $\eta$ 拓扑共轭，记作 $\xi \sim \eta$. 对于向量场来说，拓扑共轭是一过强的要求，通常使用较弱的以下概念：若有同胚 $h: M \to N$，$h$ 映 $\xi$ 的轨道为 $\eta$ 的轨道且保持轨道定向，则说 $\xi$ 与 $\eta$ 拓扑等价. 拓扑等价的向量场的轨道、奇点、闭轨等亦在同胚映射下互相对应，因此亦可认为有相同的轨道结构. 对于 $\xi$ 与 $\eta$ "局部拓扑共轭"与"局部拓扑等价"的概念，可作类似刻划.

动力系统理论的基本问题之一，是判定"充分接近"的系统是否拓扑等价. 为此需要对系统的"充分接近"赋予适当含义. 若 $f, g: V \subset X \mapsto X$ 为 $C^r$ 映射，则当 $|f^{(k)} - g^{(k)}|_0 (0 \leqslant k \leqslant r)$ 皆充分小时说 $f$ 与 $g$ 充分"$C^r$ 接近"；由此引出空间 $C^r(V, X)$ 中的"$C^r$ 拓扑". 进而通过图册过渡到 $\mathrm{Diff}^r(M)$ 上的 $C^r$ 拓扑；"$C^r$ 邻域"的含义自明，准确的叙述参考[209]. 为行文简便，本章将采用朴素的说法，以 $f$ 的"$C^r$ 近似"表示 $f$ 的某个"充分小 $C^r$ 邻域"中的任何函数. 对向量场的 $C^r$ 近似可作类似理解.

**8.1.6 定义** 设 $f \in \mathrm{Diff}^r(M)$. 若 $f$ 的每个 $C^r$ 近似拓扑共轭于 $f$，则说 $f$ 为 $C^r$ 结构稳定. 若 $f: U \subset M \to M$ 在 $a \in U$ 处为局部 $C^r$ 同胚，对 $f$ 的每个 $C^r$ 近似 $g$，$\exists\ b \in U: (f, a) \sim (g, b)$，则说 $f$ 在 $a$ 处局部 $C^r$ 结构稳定. 向量场的结构稳定性仿此定义（但用拓扑等价替代拓扑共轭）.

若 $f, g \in \mathrm{Diff}^r(M)$，$a, b \in M$，$f(a) \neq a$，$g(b) \neq b$，则可验证 $(f, a) \sim (g, b)$. 可见仅在不动点（对向量场是奇点）处才有考虑局部稳定性之必要.

主要由于 Smale 等人的大力提倡，逐渐形成了关于动力系统理论基本问题的一致看法，概括起来就是：对一定的子集 $F \subset \mathrm{Diff}^r(M)$，（Ⅰ）阐明是否每个 $f \in F$ 为（整体或局部）$C^r$ 结构稳定（稳定性问题）；（Ⅱ）探明 $F$ 是否含足够多的元素（一般性问题）；（Ⅲ）对 $F$ 中的元素依拓扑共轭进行分类（拓扑分类问题）. 对向量场可类似考虑. 以上问题迄今仅对某些特殊情况获得了令人满意的解答.

本节所述的概念容许某些自然的扩充. 首先,对任何拓扑空间 $M$ 与 $f\in\text{Homeo}(M)$, $f$ 生成一"拓扑动力系统" $\{f^n:n\in\mathbf{Z}\}$,本节对 $f\in\text{Diff}^r(M)$ 定义的种种概念依然可用. 其次,对于 $f\in C(M,M)$,可考虑"半动力系统" $\{f^n:n\in\mathbf{Z}_+\}$. 关于动力系统的种种概念经某些显然的修改之后亦可用于半动力系统.

## §2 双曲不动点

对于 $f\in\text{Diff}^r(M)$ 与 $M$ 上的 $C^r$ 向量场 $\xi$,首先考虑 $f$ 或 $\xi$ 在给定点 $x$ 处的局部行为. 若 $f(x)\neq x[\xi(x)\neq 0]$,则可指明 $f[\xi]$ 在 $x$ 附近的轨道有如一族平行移动(参考 8.5.1),这种简单形态自然不必考虑. 若 $x$ 是 $f$ 的不动点[$\xi$ 的奇点],则自然试图通过 $f$ 或 $\xi$ 在 $x$ 处的"线性化"来了解系统的局部轨道结构. 为此,给出以下定义:

**8.2.1 定义** 令 $H(X)=\{A\in\text{GL}(X):\sigma(A)\cap S^1=\varnothing\}$, $\mathscr{H}(X)=\{A\in L(X):e^A\in H(X)\}$. 分别称 $H(X)$ 与 $\mathscr{H}(X)$ 中的元为 $X$ 上的双曲自同构与双曲线性向量场. 若 $x$ 是 $C^r$ 映射 $f:U\subset M\to M$ 的不动点, $1\bar{\in}\sigma(df(x))$,则称 $x$ 为初等不动点;当 $df(x)\in H(T_xM)$ 时称 $x$ 为双曲不动点. 若 $x$ 是 $U\subset M$ 上的 $C^r$ 向量场 $\xi$ 的奇点, $\xi'(x)=\dfrac{\mathrm{d}}{\mathrm{d}t}(d\xi_t(x))|_{t=0}$,则当 $\xi'(x)\in\text{GL}(T_xM)$ 时称 $x$ 为简单奇点;当 $\xi'(x)\in\mathscr{H}(T_xM)$ 时称 $x$ 为双曲奇点.

首先指明,双曲自同构与双曲线性向量场有很简单的轨道结构.

**8.2.2 定理** 设 $A\in H(X)$,则有分解 $X=X^s\oplus X^u$,使当 $0\neq x_s\in X^s$, $0\neq x_u\in X^u$, $n\to\infty$ 时 $A^nx_s\to 0$, $A^{-n}x_u\to 0$, $|A^{-n}x_s|\to\infty$, $|A^nx_u|\to\infty$.

**注** 如通常一样,涉及谱概念时总假定相关的空间与算子已经变化.

**证** 令 $\sigma_s=\{\lambda\in\sigma(A):|\lambda|<1\}$, $\sigma_u=\sigma(A)\backslash\sigma_s$. 由标准的谱分解

结论(如见[75;4.6.6]),有分解 $X=X^s\oplus X^u$ 与 $A=A_s\oplus A_u$,使得 $A_s\in GL(X^s)$,$A_u\in GL(X^u)$,$\sigma(A_s)=\sigma_s$,$\sigma(A_u)=\sigma_u$. $\forall x\in X$,记 $x=x_s+x_u$,$x_s\in X^s$,$x_u\in X^u$. 取定 $\rho$: $1>\rho>\max\{r_\sigma(A_s),r_\sigma(A_u^{-1})\}$. 定义

$$\begin{cases}\|x_s\| = \sum\limits_{n\geqslant 0}\rho^{-n}|A^n x_s|, \ \|x_u\| = \sum\limits_{n\geqslant 0}\rho^{-n}|A^{-n}x_u|; \\ \|x\| = \max\{\|x_s\|, \|x_u\|\}.\end{cases} \tag{1}$$

显然 $|x_s|\leqslant\|x_s\|\leqslant|x_s|\sum\rho^{-n}|A_s^n|$;对 $|x_u|$ 有类似不等式,因此(1)定义 $X$ 上一等价范数 $\|\cdot\|$. 易验知 $\|Ax_s\|\leqslant\rho\|x_s\|$,因此 $\|A_s\|\leqslant\rho$;类似地 $\|A_u^{-1}\|\leqslant\rho$,于是

$$\begin{cases}\rho^{-n}\|A^n x_s\| \leqslant \|x_s\| \leqslant \rho^n\|A^{-n}x_s\|; \\ \rho^{-n}\|A^{-n}x_u\| \leqslant \|x_u\| \leqslant \rho^n\|A^n x_u\|.\end{cases} \tag{2}$$

(2)直接推出定理结论. □

称(1)所定义的 $\|\cdot\|$ 为 $A$ 的"容许范数";称 $\tau\triangleq\max\{\|A_s\|, \|A_u^{-1}\|\}$ 为 $A$ 的斜度,必 $\tau<1$. 若 $0\neq x_s\in X^s$,$0\neq x_u\in X^u$,则由(2)看出 $\|A^n x_s\|$ 与 $\|A^{-n}x_u\|$ 对 $n$ 严格下降.

完全平行于 8.2.2,有以下向量场结果:

**8.2.3 定理** 设 $A\in\mathscr{H}(X)$,则有分解 $X=X^s\oplus X^u$,使当 $0\neq x_s\in X^s$,$0\neq x_u\in X^u$,$t\to\infty$ 时 $A_t x_s\to 0$,$A_{-t}x_u\to 0$,$|A_{-t}x_s|\to\infty$,$|A_t x_u|\to\infty$,$A_t=e^{tA}$.

**证** 用 8.2.2 于 $e^A$ 得分解 $X=X^s\oplus X^u$,$e^A=B_s\oplus B_u$. 取充分小的 $b>0$,类似于(1)定义

$$\begin{cases}\|x_s\| = \int_0^\infty e^{bt}|A_t x_s|\,\mathrm{d}t, \ \|x_u\| = \int_0^\infty e^{bt}|A_{-t}x_u|\,\mathrm{d}t; \\ \|x\| = \max\{\|x_s\|, \|x_u\|\},\end{cases} \tag{3}$$

则

$$|x_s|\sum_{n\geqslant 0}e^{bn}|B_s^{-n}|^{-1}\int_0^1 e^{bt}|A_{-t}|^{-1}\mathrm{d}t\leqslant\|x_s\|$$

$$\leqslant|x_s|\sum_{n\geqslant 0}e^{bn}|B_s^n|\int_0^1 e^{bt}|A_t|\mathrm{d}t,$$

对$|x_u|$亦有类似不等式，因此(3)定义出等价范数$\|\cdot\|$. 由(3)得出$\|A_tx_s\|\leqslant e^{-bt}\|x_s\|$, $\|A_{-t}x_u\|\leqslant e^{-bt}\|x_u\|$，于是

$$\begin{cases}e^{bt}\|A_tx_s\|\leqslant\|x_s\|\leqslant e^{-bt}\|A_{-t}x_s\|;\\ e^{bt}\|A_{-t}x_u\|\leqslant\|x_u\|\leqslant e^{-bt}\|A_tx_u\|\end{cases}\tag{4}$$

(与(2)对照). (4)直接推出定理结论. □

注意若$0\neq x_s\in X^s$, $0\neq x_u\in X^u$，则(3)推出$\|A_tx_s\|$与$\|A_{-t}x_u\|$关于$t$严格下降.

8.2.2与8.2.3中的$X^s$与$X^u$分别称为$A$的稳定子空间与不稳定子空间，概称不变子空间. 若$X^s=X[X^u=X]$，则称0为$A$的吸点[斥点]；吸点显然是渐近稳定的. 若$0\neq X^s\neq X$，则称0为$A$的鞍点. 称$\dim X^s$为$A$的稳定指标；若$A\in\mathscr{H}(\mathrm{R}^n)$，则可指明$\dim X^s$等于$A$在左半平面的特征值之个数(计算重数).

现在的问题是，非线性系统在其双曲不动点附近的轨道结构是否类似于其线性化的轨道结构？这将由本节的主要结果8.2.5作出肯定回答. 证明8.2.5需要以下的

**8.2.4 引理** 设$T\in GL(X)$, $\varphi:X\to X$, $\mathrm{Lip}\varphi<|T^{-1}|^{-1}$，则$T+\varphi\in\mathrm{Homeo}(X)$.

**证** $\forall\ y\in X$，令$F_y=T^{-1}(y-\varphi)$，则$\mathrm{Lip}F_y\leqslant|T^{-1}|.\mathrm{Lip}\varphi<1$，于是有唯一$x\in X: x=F_yx$，即$Tx+\varphi(x)=y$. 可见$T+\varphi:X\to X$为双射. 若$Tx_i+\varphi(x_i)=y_i$，则

$$\begin{aligned}|y_1-y_2|&\geqslant|Tx_1-Tx_2|-|x_1-x_2|\mathrm{Lip}\varphi\\&\geqslant(|T^{-1}|^{-1}-\mathrm{Lip}\varphi)|x_1-x_2|,\end{aligned}$$

这表明$(T+\varphi)^{-1}$连续，因此$T+\varphi\in\mathrm{Homeo}(X)$. □

**8.2.5 定理** (Hartman, 1964)设$f=A+\varphi:\Omega\subset X\to X$为$C^{-1}$映射，$f(0)=0\in\Omega$, $A=f'(0)\in H(X)$，则$(f,0)\sim(A,0)$.

**证** 可设$\mathrm{Lip}\varphi$充分小且$\varphi\in Z\triangleq C_b(X,X)=\{\psi\in C(X,X):|\psi|_0<\infty\}$(否则取充分小的$\delta>0$，在$B_\delta(0)$外以$\varphi(\delta x/|x|)$代$\varphi(x)$). 由8.2.4有$f\in\mathrm{Homeo}(X)$. 为证有$h\in\mathrm{Homeo}(X)$使$fh=hA$

(8.1.5)，只需证以下命题：

(P)　若 $\varphi$、$\psi\in Z$，Lip$\psi$ 与 Lip$\psi$ 充分小，则存在唯一

$$h=I+\eta:\eta'\in Z,(A+\psi)h=h(A+\varphi).$$

事实上，若(P)成立，则有 $h_i=I+\eta_i,\eta_i\in Z(i=1,2)$，使 $(A+\varphi)h_1=h_1A,Ah_2=h_2(A+\varphi)$. 这推出 $(A+\varphi)h_1h_2=h_1h_2(A+\varphi),h_1h_2-I\in Z$，于是由唯一性有 $h_1h_2=I$；同理 $h_2h_1=I$，于是 $h=h_1$ 合于定理所求. 下面证命题(P).

设 $X^s,X^u,A_s,A_u$，$\|\cdot\|$ 依 8.2.2，$P_s:X\to X^s$ 记投影，令 $\eta_s=P_s\eta$ $(\eta\in Z)$，$Z^s=C_b(X,X^s)$；$\varphi_s,\psi_s,\eta_u,Z^u,\varphi_u,\psi_u$ 仿此. 在 $Z=Z^s\oplus Z^u$ 中采用范数 $\|\eta\|_0=\sup\|\eta(x)\|$. 由

$$\begin{aligned}F\eta=&((A_s\eta_s+\psi_s(I+\eta)-\varphi_s)(A+\varphi)^{-1},\\&A_u^{-1}(\eta_u(A+\varphi)-\psi_u(I+\eta)+\varphi_u))\end{aligned}\tag{5}$$

定义出 $F:Z\to Z$. 因

$$\text{Lip}F\leqslant\max\{\|A_s\|+\text{Lip}\psi,\|A_u^{-1}\|(1+\text{Lip}\psi)\}<1,$$

故有唯一 $\eta\in Z:\eta=F\eta$. $h=I+\eta$ 即合命题(P)之要求. □

平行于 8.2.5 的向量场结果是：

**8.2.6 定理** (Hartman-Grobman) 设 $\xi=A+\varphi:\Omega\subset X\to X$ 是 $C^1$ 向量场，$\xi(0)=0$，$A=\xi'(0)\in\mathscr{H}(X)$，则 $\xi$ 与 $A$ 的流在 $x=0$ 处局部拓扑共轭.

**证**　如同 8.2.5 之证，可设 $\varphi\in Z\triangleq C_b(X,X)$，$\varepsilon\triangleq\text{Lip}\varphi$ 充分小. 由标准的微分方程结论知 $\xi$ 的流定义在 $\mathbf{R}\times X$ 上，且 $\text{Lip}\xi_t\leqslant e^{\mu t}$，$\mu=|A|+\varepsilon$. 令 $A_t=e^{tA}$，$\theta_t=\xi_t-A_t$. $\forall\ x\in X,t>0$，有

$$|\theta_tx|\leqslant\int_0^t|A\theta_\tau x+\varphi(\xi_\tau x)|\mathrm{d}\tau\leqslant t|\varphi|_0+|A|\int_0^t|\theta_\tau x|\mathrm{d}\tau,$$

于是由 Gronwall 不等式得 $|\theta_tx|\leqslant|t\varphi|_0e^{|tA|}$. 显然此不等式亦适用于 $t\leqslant0$，因此 $\theta_t\in Z(\forall\ t\in\mathbf{R})$. $\forall\ t\in J$，由

$$|\theta_tx-\theta_ty|\leqslant\int_0^t|A(\theta_\tau x-\theta_\tau y)+\varphi(\xi_\tau x)-\varphi(\xi_\tau y)|\mathrm{d}\tau$$

$$\leqslant \int_0^t (|A|\,|\theta_\tau x - \theta_\tau y| + \varepsilon e^{\mu\tau}|x-y|)\mathrm{d}\tau$$

$$\leqslant \varepsilon e^{\mu}|x-y| + |A|\int_0^t |\theta_\tau x - \theta_\tau y|\mathrm{d}\tau$$

得 $|\theta_1 x - \theta_1 y| \leqslant \varepsilon e^{|A|+\mu}|x-y|$，因此 $\mathrm{Lip}\theta_1 \leqslant \varepsilon e^{|A|+\mu}$. 对 $A_1+\theta_1$ 应用 8.2.5 得唯一 $h = I+\eta \in \mathrm{Homeo}(X)$：$\eta \in Z$，$hA_1 = \xi_1 h$. $\forall\, t\in \mathrm{R}$，由 $\xi_t h A_{-t} A_1 = \xi_1 \xi_t h A_{-t}$，$\xi_t h A_{-t} - I = A_t \eta A_{-t} + \theta_t h A_{-t} \in Z$ 及唯一性得 $\xi_t h A_{-t} = h$，即 $\xi_t h = h A_t$. □

**8.2.7 推论** 设 $A\in H(X)$，$B\in L(X)$，$|A-B|$ 充分小，则 $A \overset{\mathrm{loc}}{\sim} B$；因此 $\{B : A\overset{\mathrm{loc}}{\sim} B\} \subset L(X)$ 为开集.

**证** 取 $u\in C_c^1(\mathrm{R})$，使在 $t=0$ 邻近 $u(t)=1$. 令 $\varphi(x) = u(|x|)(Bx - Ax)$，则 $\varphi \in Z$，$\mathrm{Lip}\varphi \leqslant \mathrm{const}\,|A-B|$，在 $x=0$ 邻近 $(A+\varphi)x = Bx$，于是由 8.2.5 推出所要结论. □

**8.2.8 引理** 设 $A\in L(X)$. 则 $\forall\, \varepsilon>0$，$\exists\, \delta>0$，$\forall\, B\in L(X)$，当 $|A-B|<\delta$ 时 $\sigma(B)\subset \sigma(A)+D_\varepsilon(0)$，$D_\varepsilon(0) = \{\lambda\in \mathrm{C} : |\lambda|<\varepsilon\}$.

**证** 若引理不真，则有 $\varepsilon>0$，$B_n\in L(X)$，使 $|A-B_n|\to 0$，且有 $\lambda_n \in \sigma(B_n)\setminus[\sigma(A)+D_\varepsilon(0)]$. 可设 $\lambda_n\to\lambda$，必定 $\lambda \overline{\in} \sigma(A)$，于是 $\lambda I - A \in \mathrm{GL}(X)$. 这推出对大的 $n$ 有 $\lambda_n I - B_n \in \mathrm{GL}(X)$，与 $\lambda_n \in \sigma(B_n)$ 矛盾. □

**8.2.9 引理** 设 $x_0$ 是 $C^r$ 映射 $f : U\subset M\to M$ 的初等不动点，则 $f$ 的每个 $C^r$ 近似 $g$ 在 $x_0$ 邻近有唯一不动点 $x_g$，且 $x_g$ 是初等不动点；当 $x_0$ 是 $f$ 的双曲不动点时 $x_g$ 亦是 $g$ 的双曲不动点.

对向量场的奇点有类似结论.

**证** 因问题是局部的，不妨设 $f : U\subset X\to X$，$f(0)=0\in U$，$1\overline{\in}\sigma(f'(0))$. 于是对映射

$$C^r(U,X)\times U \to X,\ (g,x)\mapsto x - g(x)$$

应用隐函数定理并结合 8.2.8 得所要证. □

**8.2.10 定理** 设 $C^r$ 映射 $f : U\subset M\to M$ 以 $x_0$ 为双曲不动点，则 $f$ 在 $x_0$ 处局部 $C^r$ 结构稳定.

对向量场有类似结论.

**证** 设 $g$ 是 $f$ 的 $C^r$ 近似. 由 8.2.9, $g$ 在 $x_0$ 邻近有唯一不动点 $x_g$ 且 $x_g$ 是双曲不动点. 由 8.2.5, 有 $f\overset{\text{loc}}{\sim}df(x_0)$, $g\overset{\text{loc}}{\sim}dg(x_g)$; 由 8.2.7, 有 $df(x_0)\overset{\text{loc}}{\sim}dg(x_g)$, 于是 $(f,x_0)\sim(g,x_g)$, 即 $f$ 在 $x_0$ 处局部 $C^r$ 结构稳定.

## §3 拓扑分类·一般性

本节给出 $H(\mathrm{R}^n)$ 与 $\mathscr{H}(\mathrm{R}^n)$ 的拓扑分类. 由 8.2.5 与 8.2.6, 这相当于对 $n$ 维流形上的 $C^r$ 映射的双曲不动点与 $C^r$ 向量场的双曲奇点作出了拓扑分类.

**8.3.1 引理** 设 $A,B\in H(X)$, $(A,0)\sim(B,0)$, 则 $A\sim B$.

**证** 设 $A$ 与 $B$ 的不变子空间分别为 $X^s,X^u$ 与 $Y^s,Y^u$; $U,V$ 是 $X$ 的 0-邻域, 同胚 $h:U\cup AU\to V\cup BV$ 满足 $hU=V$, $h(0)=0$, $hA|U=Bh|U$. 令 $U^s=U\cap X^s$, $V^s=V\cap Y^s$; $U^u$, $V^u$ 仿此. 不妨设 $U=\{x:\|x\|<\delta\}$, $\|\cdot\|$ 是 $A$ 的容许范数. $\forall\, x\in U^s$, 必 $A^kx\in U\,(k\geqslant 0)$, $A^kx\to 0\,(k\to\infty)$, 于是 $B^khx=hA^kx\to 0$, 因而 $hx\in V^s$, 故 $hU^s\subset V^s$. 同理 $hU^s\supset V^s$, 因此 $hU^s=V^s$. 类似地 $hU^u=V^u$. $\forall\, x_s\in X^s$, 取 $k>0$, 使 $A^kx_s\in U^s$; 定义 $h_s(x_s)=B^{-k}hA^kx_s$, 则 $h_s(x_s)$ 与 $k$ 无关, 且可验知 $h_s:X^s\to Y^s$ 为同胚. 类似地定义 $h_u:X^u\to Y^u$, 则 $h\triangleq h_s\times h_u\in\mathrm{Homeo}(X)$ 满足 $hA=Bh$, 因此 $A\sim B$. □

**8.3.2 引理** 若 $\beta\subset H(X)$ 是一连通子集, $A,B\in\beta$, 则 $A\sim B$.

**证** 由 8.2.7 与 8.3.1, $\beta$ 中每个"$\sim$等价类"是相对开集. 由连通性, $\beta$ 只能含一个$\sim$等价类. □

**8.3.3 定理** 任何 $A\in H(\mathrm{R}^n)$ 拓扑共轭于以下 $4n$ 个互不共轭的标准形之一:

$$\mathrm{diag}\Big(\pm\frac{1}{2},\frac{1}{2},\cdots,\frac{1}{2},\overbrace{\pm 2,2,\cdots,2}^{i}\Big),\ i=0,1,\cdots,n. \qquad (1)$$

**证** 首先证 $A\in H(\mathrm{R}^n)$ 必共轭于(1)中某个标准形，为此令 $A$ 沿 $H(\mathrm{R}^n)$ 内的单参数曲线作连续变形(用 8.3.2)，直至化为标准形. 这非常类似于 1.7.5 之证，因此只需述其大意. 不妨设 $A$ 是 Jordan 矩阵：$A=J_1\oplus\cdots\oplus J_r$，$J_k(1\leqslant k\leqslant r)$ 是以下形式之一：

$$\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1\\ & & & \lambda\end{pmatrix},\quad \begin{pmatrix}\alpha & -\beta & 1 & & & & \\ \beta & \alpha & 0 & 1 & & & \\ & & \alpha & -\beta & 1 & & \\ & & \beta & \alpha & 0 & 1 & \\ & & & & \ddots & \ddots & \ddots\end{pmatrix},\tag{2}$$

其中 $\lambda,\alpha,\beta\in\mathrm{R}$，$|\lambda|\neq 0,1$，$\alpha^2+\beta^2\neq 0,1$. 将(2)中的 1 换成参数 $t$ 并令其连续地降至零. 其次设 $\alpha+i\beta=\rho e^{i\theta}$，将 $\theta$ 换成参数 $s$ 并令其连续地变到零. 这样，$A\sim\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，$|\lambda_j|\neq 0,1(1\leqslant j\leqslant n)$. 再经几个明显的"连续变形"步骤，即得出 $A$ 共轭于(1)中某个标准形.

其次证(1)中各标准形互不共轭. 设 $A,B$ 是(1)中的两个标准形. 若 $A\sim B$，则有 $h\in\mathrm{Homeo}(R^n)$：$hA=Bh$. 设 $A$ 与 $B$ 的稳定子空间分别为 $X^s$ 与 $Y^s$，则必 $X^s\cong Y^s$，从而 $A,B$ 中所含 $\pm1/2$ 之个数相同. 利用 $hA=Bh$，由一个度论论证可指明 $-1/2$ 在 $A,B$ 中同时出现或不出现. 对 $-2$ 亦有同样结论，因此 $A=B$. □

与 8.3.3 相对应，对 $\mathscr{H}(\mathrm{R}^n)$ 有以下分类定理：

**8.3.4 定理** 任何 $A\in\mathscr{H}(\mathrm{R}^n)$ 拓扑共轭于以下 $n+1$ 个互不共轭的标准形之一：

$$\mathrm{diag}(-1,\cdots,-1,\overbrace{1,\cdots,1}^{i\text{个}}),\ i=0,1,\cdots,n.\tag{3}$$

**证** 只需指明：任给 $A,B\in\mathscr{H}(\mathrm{R}^n)$，$A\sim B\Leftrightarrow A$ 与 $B$ 的稳定指标相同. 设 $A$ 与 $B$ 的不变子空间分别为 $X^s,X^u$ 与 $Y^s,Y^u$. 当 $A\sim B$ 时显然 $\dim X^s=\dim Y^s$. 反之，若 $\dim X^s=\dim Y^s$，则有线性同构 $\varphi_s: X^s\to Y^s$，$\varphi_u: X^u\to Y^u$，今证 $A\sim B$. 设 $A_t=e^{tA}$，$B_t=e^{tB}$. 若有 $h_s\in\mathrm{Homeo}(Y^s)$，使 $h_s\varphi_sA_t\varphi_s^{-1}\equiv B_th_s$，$h_u$ 仿此，则 $h=h_s\varphi_s\times h_u\varphi_u\in\mathrm{Homeo}(\mathrm{R}^n)$，$hA_t\equiv$

$B_t h$，从而 $A \sim B$. 不妨只构作如上的 $h_s$($h_u$ 的构成是类似的). 以 $X^s$ 代 $R^n$，又不妨设 $X^s = Y^s = R^n$，$\varphi_3 = \text{id}$. 于是要求 $h \in \text{Homeo}(R^n)$ 使 $hA_t = B_t h$. 分别以 $\| \cdot \|$ 与 $\| \cdot \|'$ 记 $R^n$ 上关于 $A, B$ 的容许范数，$S$ 与 $S'$ 是相应的单位球面. 令 $h(x) = x / \| x \|'$，则 $h: S \to S'$ 为同胚. 扩张 $h$ 如下：令 $h(0) = 0$；若 $0 \neq x \in R^n$，则 $\| A_t x \|$ 对 $t$ 严格下降，故有唯一 $\tau \in R$：$A_\tau x \in S$，定义 $h(x) = B_{-\tau} h A_\tau x$，如此得到 $h: R^n \to R^n$. $h$ 显然是双射且 $hA_t = B_t h$，只需证 $h$ 连续(同理得出 $h^{-1}$ 连续). 设 $x_k \to x (k \to \infty)$，$A_{\tau_k} x_k \in S$. 若 $x \neq 0$，$A_\tau x \in S$，则必 $\tau_k \to \tau$，于是 $h(x_k) \to h(x)$. 若 $x = 0$，则必 $\tau_k \to -\infty$，$h(x_k) \to 0$. 因此 $h$ 连续. □

结合 8.3.3，8.3.4 与 8.2.5，8.2.6 得出结论：若 $\dim M = n$，则 $M$ 上的 $C^r$ 映射在其双曲不动点处的局部行为只有 $4n$ 种不同类型；或者说只有 $4n$ 种拓扑上不同的双曲不动点，其中每种可用一个标准的线性变换作为说明. 对 $M$ 上的 $C^r$ 向量场的双曲奇点有类似结论.

至此可以说，对双曲不动点[奇点]的局部性质已有了完全的了解. 于是自然提出问题：在多大的程度上有理由假定动力系统仅有双曲不动点[奇点]？这就进入"一般性"的课题. 首先可以说，有限维线性系统"一般"是双曲的，严格地说是：

**8.3.5 命题** $H(R^n)$ 与 $\mathscr{H}(R^n)$ 分别是 $GL(R^n)$ 与 $L(R^n)$ 的稠密开子集.

**证** 直接由 8.2.8 推出 $H(R^n)$ 在 $GL(R^n)$ 中开，从而 $\mathscr{H}(R^n) = \exp^{-1} H(R^n)$ 在 $L(R^n)$ 中开. 任给 $A \in GL(R^n)$，当 $\rho > 1$ 充分接近 1 时 $\rho A \in H(R^n)$(否则有 $\rho_k \downarrow 1$，$\lambda_k \in \sigma(\rho_k A) \cap S^1$，这推出 $\sigma(A)$ 含一无限集 $\{\lambda_k / \rho_k\}$！)，这表明 $H(R^n)$ 在 $GL(R^n)$ 中稠. 类似地可证 $\mathscr{H}(R^n)$ 在 $L(R^n)$ 中稠. □

为以上结果所鼓舞，似乎可推断 $C^r$ 同胚"一般"仅有双曲不动点. 在一定条件下确实如此. 为达此结论，需要某些拓扑准备. 设 $f \in C^1(M, N)$，$Q$ 是 $N$ 的子流形. 若当 $x \in M$，$y = f(x) \in Q$ 时恒有 $df_x(T_x M) + T_y Q = T_y N$，则说 $f$ 横截 $Q$.

**8.3.6 引理** 设 $M$ 是 $m$ 维 $C^r$ 紧流形，$N$ 是 $n$ 维 $C^r$ 流形，$Q$ 是 $N$ 的 $C^r$ 子流形，则 $\{f\in C^r(M,N): f$ 横截 $Q\}$ 是 $C^r(M,N)$ 中的稠开集.

关于以上结果可参看[92].

**8.3.7 定理** 设 $M$ 是 $n$ 维 $C^r$ 紧流形，$G_0=\{f\in \mathrm{Diff}^r(M): f$ 只有初等不动点$\}$，$G_1=\{f\in G_0: f$ 只有双曲不动点$\}$，则 $G_0, G_1$ 是 $\mathrm{Diff}^r(M)$ 中的稠开集.

**证** 设 $f\in \mathrm{Diff}^r(M)$，$x\in M$. 则 $x$ 是 $f$ 的初等不动点 $\Leftrightarrow f(x)=x$ 且 $\mathrm{id}-df(x)\in GL(T_xM)\Leftrightarrow(\mathrm{id},f)x\in\Delta$ 且 $(\mathrm{id},df(x))T_xM+T_{(x,x)}\Delta=T_xM\times T_xM$，此处 $\Delta=\{(x,x): x\in M\}$. 因此 $G_0=\{f\in \mathrm{Diff}^r(M): (\mathrm{id},f)$ 横线 $\Delta\}$，于是由 8.3.7 得出 $G_0$ 是 $\mathrm{Diff}^r(M)$ 中的稠开集.

其次证 $G_1$ 在 $\mathrm{Diff}^r(M)$ 中开. 取定 $f\in G_1$. 由 8.2.9，$f$ 只有孤立不动点，因此 $\mathrm{Fix}f=\{x_i\}$ 是有限集. 若 $g\in \mathrm{Diff}^r(M)$ 是 $f$ 的 $C^r$ 近似，则 $g$ 仅在每个 $x_i$ 邻近有唯一不动点且是双曲不动点(8.2.9)，因此 $g\in G_1$. 这表明 $G_1$ 是开集.

最后证 $G_1$ 在 $G_0$ 中稠(从而在 $\mathrm{Diff}^r(M)$ 中稠). 取定 $f\in G_0$. 如同上段，由 8.2.9 推出 $\mathrm{Fix}f=\{x_i\}$ 是有限集，$f$ 的任何 $C^r$ 近似 $g$ 仅有初等不动点 $y_i$，$y_i\in U_i$，$U_i$ 是 $x_i$ 的充分小的邻域. 于是只需指明：在每个 $U_i$ 内对 $f$ 作微小修正可使 $x_i$ 成为双曲不动点. 所需局部修正可在 $\mathbb{R}^n$ 上描述：设 $f: U\subset\mathbb{R}^n\to\mathbb{R}^n$ 以 $0\in U$ 为唯一不动点，$f'(0)\in GL(\mathbb{R}^n)$. 取 $\varphi\in C_c^r(\mathbb{R}^n)$，使在 $x=0$ 邻近 $\varphi(x)=1$；令 $g(x)=f((1+\varepsilon\varphi(x))x)$，$\varepsilon>0$ 充分小，则 $g$ 是 $f$ 的 $C^r$ 近似，在 $\mathbb{R}^n$ 的某个 0-邻域外 $g(x)=f(x)$，$g'(0)=(1+\varepsilon)f'(0)\in H(\mathbb{R}^n)$(参考 8.3.5 之证)，可见 $g$ 以 0 为双曲不动点. □

**注** 用类似的方法可对向量场建立与 8.3.7 相应的结果.

## §4 稳定流形

若 $x_0$ 是 $f\in \mathrm{Diff}^r(M)$ 的不动点，则依 8.1.4 有

$$W^s(x_0)=\{x\in M: f^n(x)\to x_0(n\to\infty)\};\qquad (1)$$

$$W^u(x_0)=\{x\in M: f^{-n}(x)\to x_0(n\to\infty)\}.\qquad (2)$$

本节要指明，当 $x_0$ 是双曲不动点时 $W^s(x_0)$ 与 $W^u(x_0)$ 具有自然的 $C^r$ 流形结构. 对于 $A\in H(X)$，由 8.2.2 有 $W^s(0,A)=X^s, W^u=(0,A)=X^u$，$X^s,X^u$ 是 $A$ 的不变子空间. 利用这一线性结果，可以描述 $W^s(x_0)$ 与 $W^u(x_0)$ 在 $x_0$ 邻近的局部结构；由此经一适当构造即获得 $W^s(x_0)$ 与 $W^u(x_0)$ 的 $C^r$ 流形结构. 下面不妨只考虑 $W^s(x_0)$. 以下局部结果是关键的.

**8.4.1 定理** 设 $C^r$ 映射 $f:\Omega\subset X\to X$ 以 $0\in\Omega$ 为双曲不动点，$A=f'(0)$，$X^s,X^u$ 与 $\|\cdot\|$ 依 8.2.2，在 $X$ 中采用范数 $\|\cdot\|$，令 $X_\varepsilon=\{x\in X: \|x\|<\varepsilon\}$，$X_\varepsilon^s,X_\varepsilon^u$ 仿此. 若 $\varepsilon>0$ 充分小，则存在 $g\in C^r(X_\varepsilon^s,X_\varepsilon^u)$，使得 $g(0)=0, g'(0)=0, \mathrm{Gr}g=W_\varepsilon^s(0)\triangleq\{x: O^+(x,f)\subset X_\varepsilon\}$.

**证** 令 $\varphi=f-A$. 因问题是局部的，可设 $\delta\triangleq\mathrm{Lip}\varphi$ 充分小(参考 8.2.5 之证). 设 $A_s,A_u$ 依 8.2.2 之证，$P_s:X\to X^s$ 与 $P_u:X\to X^u$ 是投影，$f_s=p_sf$，$f_u,\varphi_s,\varphi_u$ 等仿此. 令 $\Sigma=\{\sigma: Z_+\to X\,|\,\sigma(k)\to 0(k\to\infty)\}$，则 $\Sigma$ 依 sup 范数 $\|\cdot\|_0$ 是一Banach 空间. 取适当小的 $\rho>0$，令 $\Sigma_\rho=\{\sigma\in\Sigma: \|\sigma\|_0<\rho\}$. 由

$$F(x_s,\sigma)(k)=\begin{cases}(x_s,A_u^{-1}(\sigma_u(1)-\varphi_u(\sigma(0)))), & k=0;\\ (f_s(\sigma(k-1)),A_u^{-1}(\sigma_u(k+1)-\varphi_u(\sigma(k)))), & k\geqslant 1\end{cases}\qquad (3)$$

定义一映射 $F: X_\rho^s\times\Sigma_\rho\to\Sigma$. 直接看出 $F\in C^r$，且经某些计算后得出

$$F_{x_s}(x_s,\sigma)\xi=(\xi,0,0,\cdots),(x_s,\sigma,\xi)\in X_\rho^s\times\Sigma_\rho\times X^s;\qquad (4)$$

$$F_\sigma(0,0)(\beta)(k)=\begin{cases}(0,A_u^{-1}\beta_u(1)), & k=0;\\ (A_s\beta_s(k-1),A_u^{-1}\beta_u(k+1)), & k\geqslant 1,\end{cases} \tag{5}$$

(5)中 $\beta\in\Sigma$. 由(5)看出 $\|F_\sigma(0,0)\|\leqslant\tau\triangleq\max\{\|A_s\|,\|A_u^{-1}\|\}<1$. 由隐函数定理,有 $\varepsilon\in(0,\rho)$,从方程 $F(x_s,\sigma)=\sigma$ 可解出 $C^r$ 函数 $\sigma=\sigma(x_s):X_\varepsilon^s\to\Sigma_\rho$, $\sigma(0)=0$. 令 $g(x_s)=\sigma_u(x_s)(0)$,则 $g\in C^r(X_\varepsilon^s,X_\rho^u)$, $g(0)=0$. 下面验明 $g$ 合于定理要求.

1°证 $g'(0)=0$. 由 $F(x_s,\sigma(x_s))=\sigma(x_s)$得

$$F_{x_s}(0,0)\xi+F_\sigma(0,0)\sigma'(0)\xi=\sigma'(0)\xi, \tag{6}$$

其中 $\xi\in X^s$. 结合(4)(5)(6)得出

$$(\sigma'_s(0)\xi)(k)=A_s(\sigma'_s(0)\xi)(k-1)=\cdots=A_s^k\xi; \tag{7}$$

$$\begin{aligned}(\sigma'_u(0)\xi)(k)&=A_u^{-1}(\sigma'_u(0)\xi)(k+1)=\cdots\\ &=A_u^{-n}(\sigma'_u(0)\xi)(k+n).\end{aligned} \tag{8}$$

在(8)中令 $n\to\infty$得$(\sigma'_u(0)\xi)(k)=0(\xi\in X^s,k\geqslant 0)$,因此 $\sigma'_u(0)=0$,从而 $g'(0)=o$. 必要时适当缩小 $\varepsilon$,不妨设 $gX_\varepsilon^s\subset X_\varepsilon^u$.

2°证 $\mathrm{Gr}g\subset W_\varepsilon^s(0)$. (7)推出 $\|(\sigma'_s(0)\xi)(k)\|\leqslant\tau\|\xi\|\ (k\geqslant 1)$;这结合 $\sigma'_u(o)=0$ 知对小的 $\|\xi\|$ 有 $\|\sigma(\xi)\|_0\leqslant\|\xi\|$,因此不妨设 $\sigma(\cdot):X_\varepsilon^s\to\Sigma_\varepsilon$. 由 $F(\xi,\sigma(\xi))=\sigma(\xi)$与(3)推出

$$\sigma(\xi)(k)=f(\sigma(\xi)(k-1))=\cdots=f^k(\sigma(\xi)(0))\in X_\varepsilon,$$

这表明 $0^+(\xi,g(\xi))\subset X_\varepsilon(\xi\in X_\varepsilon^s)$,因此 $\mathrm{Gr}g\subset W_\varepsilon^s(0)$.

3°建立一个辅助不等式. 设 $x=(x_s,x_u)$, $y=(y_s,y_u)$, $x,y\in X_\varepsilon$, $\|x_s-y_s\|\leqslant\|x_u-y_u\|$,则

$$\begin{aligned}&\|f_s(x)-f_s(y)\|=\|A_s(x_s-y_s)+\varphi_s(x)-\varphi_s(y)\|\\ &\leqslant(\tau+\delta)\|x-y\|\leqslant(\tau^{-1}-\delta)\|x-y\|\\ &\leqslant\|A_u(x_u-y_u)+\varphi_u(x)-\varphi_u(y)\|=\|f_u(x)-f_u(y)\|,\end{aligned}$$

以上用到 $\tau+\delta<1<\tau^{-1}-\delta$. 由此得

$$(\tau^{-1}-\delta)\|x-y\|\leqslant\|f(x)-f(y)\|=\|f_u(x)-f_u(y)\|. \tag{9}$$

4°证 $\mathrm{Gr}g \supset W_\varepsilon^s(0)$. 设 $x=(x_s,x_u)\in W_\varepsilon^s(0)$, $y=(x_s,g(x_s))$，则反复运用(9)得

$$\|x-y\| \leqslant (\tau^{-1}-\delta)^{-k}\|f^k(x)-f^k(y)\| \leqslant 2\varepsilon(\tau^{-1}-\delta)^{-k},$$

令 $k\to\infty$ 得 $x=y\in\mathrm{Gr}g$. □

在 8.4.1 的条件下，$(\mathrm{id},g):X_\varepsilon^s\to X_\varepsilon$ 是一 $C^r$ 嵌入，因此 $W_\varepsilon^s(0)=(\mathrm{id},g)X_\varepsilon^s$ 是 $X$ 的 $C^r$ 子流形，它在 $x=0$ 处以 $X^s$为切空间(7.1.4). 将这一局部结论过渡到流形上得出：若 $x_0$ 是 $f\in\mathrm{Diff}^r(M)$的双曲不动点，则有 $x_0$ 的充分小的邻域 $V$，使 $W_V^s(x_0)\triangleq\{x:O^+(x,f)\subset V\}$是 $M$ 的 $C^r$ 子流形，$W_V^s(x_0)$在 $x_0$ 以 $df(x_0)$的稳定子空间为切空间. 进而有如下整体结论：

**8.4.2 定理** 若 $x_0$ 是 $f\in\mathrm{Diff}^r(M)$的双曲不动点，则 $W^s(x_0)$是 $M$ 的 $C^r$ 浸入子流形(即所谓稳定流形).

**证** 以 $E^s$ 记 $df(x_0)$的稳定子空间，取充分小的 $\varepsilon>0$，以 $B_\varepsilon^s$ 记 $E^s$ 中以 0 为心的 $\varepsilon$-球. 由 8.4.1 知有 $M$ 的 $C^r$ 子流形 $S_0$ 及 $C^r$ 同胚 $\varphi_0:S_0\to B_\varepsilon^s$, $S_0=W_V^s(x_0)$, $V$ 是 $x_0$ 的某个邻域. 令 $S_k=f^{-k}S_0$，则 $W^s(x_0)=\bigcup_0^\infty S_k$. 令 $\varphi_k=\varphi_0 f^k(k\geqslant 0)$，则 $\varphi_k\varphi_j^{-1}=\varphi_0 f^{k-j}\varphi_0^{-1}$ 是 $C^r$ 同胚，$\varphi_k S_k=B_\varepsilon^s(k,j\geqslant 0)$，因此图册$\{(S_k,\varphi_k)\}$在 $W^s(x_0)$上导入一 $C^r$ 流形结构，使 $W^s(x_0)$成为 $M$ 的 $C^r$ 浸入子流形. □

以上讨论完全可平行地移用于 $W^u(x_0)$. 因此，若 $x_0$ 是 $f\in\mathrm{Diff}^r(M)$的双曲不动点，则 $W^u(x_0)$是 $M$ 的 $C^r$ 浸入子流形，即所谓不稳定流形；$W^u(x_0)$在 $x_0$ 以 $df(x_0)$的不稳定子空间为切空间，因此 $W^s(x_0)$与 $W^u(x_0)$在 $x_0$ 横截相交(参考 § 7.1).

现将以上结果过渡到向量场. 设 $\xi$ 是 $M$ 上的 $C^r$ 向量场，$x_0$ 是 $\xi$ 的奇点. 对应于(1)(2)有

$$W^s(x_0,\xi)=\{x\in M:\xi_t(x)\to x_0(t\to\infty)\};\tag{10}$$

$$W^u(x_0,\xi)=\{x\in M:\xi_t(x)\to x_0(t\to-\infty)\}.\tag{11}$$

由 $$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{d}\xi_t(x_0))=\frac{\mathrm{d}}{\mathrm{d}s}(\mathrm{d}\xi_s(x_0)\mathrm{d}\xi_t(x_0))|_{s=0}=\xi'(x_0)\mathrm{d}\xi_t(x_0)$$

得 $\mathrm{d}\xi_t(x_0)=\exp t\xi'(x_0)$(参考 8.2.1). 因此,若 $x_0$ 是 $\xi$ 的双曲奇点,则 $x_0$ 是 $C^r$ 同胚 $\xi_1$ 的双曲不动点. 任给 $x_0$ 的邻域 $V$,存在 $x_0$ 的邻域 $U$,使得 $\forall\ t\in J:\xi_t U\subset V$. $\forall\ x\in W^s(x_0,\xi_1)$,当 $t$ 充分大,$k\leqslant t<k+1$ 时 $\xi_t(x)=\xi_{t-k}\xi_1^k(x)\in\xi_{t-k}U\subset V$,可见 $x\in W^s(x_0,\xi)$. 其次显然 $W^s(x_0,\xi)\subset W^s(x_0,\xi_1)$,因此 $W^s(x_0,\xi)=W^s(x_0,\xi_1)$. 于是应用 8.4.2 于 $f=\xi_1$ 得

**8.4.3 定理** 设 $\xi$ 是 $M$ 上的 $C^r$ 向量场,$x_0$ 是 $\xi$ 的双曲奇点,则 $W^s(x_0,\xi)$是 $M$ 的 $C^r$ 浸入子流形,它在 $x_0$ 以 $\xi'(x_0)$的稳定子空间为切空间.

自然称 8.4.3 中的 $W^s(x_0,\xi)$为 $\xi$ 在 $x_0$ 的稳定流形. 对于"不稳定流形"$W^u(x_0,\xi)$可类似讨论.

## §5 双曲闭轨

本节设 $\xi$ 是 $M$ 上的 $C^r$ 向量场,$\Gamma$ 是 $\xi$ 的 $p$-周期闭轨,$p>0$(参考 8.1.2).

基本的问题是阐明 $\xi$ 在 $\Gamma$ 邻近的轨道结构,为此导入"Poincaré 映射"这一辅助工具. 若超曲面 $S\subset M$ 在点 $x_0\in\Gamma$ 与 $\Gamma$ 横截相交(这意味着 $\xi(x_0)\overline{\in}T_{x_0}S$),则称 $S$ 为 $\Gamma$ 在 $x_0$ 处的截面. 给定 $\Gamma$ 在 $x_0$ 的截面 $S$,由流的连续性,有 $x_0$ 在 $S$ 上的邻域 $V$,$\forall\ x\in V$,存在最小的 $t\coloneqq t(x)>0$($t$ 必定近似于 $p$),使得 $\xi_t(x)\in S$ 且 $\xi_t(x)$邻近 $x_0$. 称映射

$$P:V\to S,x\mapsto\xi_{t(x)}(x)$$

为 $\Gamma$ 在 $S$ 上(或在 $x_0$ 处)的 Poincaré 映射. 显然 $Px_0=x_0$. 本节指明,当 $x_0$ 是 $P$ 的双曲不动点时,$P$ 在 $x_0$ 邻近的局部性态完全决定了 $\xi$ 在 $\Gamma$ 邻近的轨道结构;而对于 $P$ 的局部性态的描述已有前面几节的结果可用.

首先建立几个预备结果.

**8.5.1 直化引理** 若 $\xi(x_0)\neq 0$,则有 $M$ 在 $x_0$ 处的图$(U,\varphi)$与 0

$\neq e\in X$，使得$\forall\ x\in U$：$d\varphi_x\xi(x)=e$.

**证** 取 $M$ 在 $x_0$ 处的图$(V,\psi)$，使 $\psi(x_0)=0$. 必定 $e=d\psi_{x_0}\xi(x_0)\neq 0$. 设 $X=Re\oplus Y$，定义

$$f(te+y)=\psi\xi_t\psi^{-1}(y),t\in R,y\in Y,$$

则 $f$ 是 $X$ 的某个 0-邻域上的 $C^r$ 映射，$f(0)=0$. 由

$$\frac{d}{dt}f(te)|_{t=0}=d\psi_{x_0}\frac{d}{dt}\xi_t(x_0)|_{t=0}=e$$

与 $f(y)=y(\forall\ y\in Y)$得出 $f'(0)=id$，因此 $f$ 在 $x=0$ 处为局部 $C^r$ 同胚. 取 $x_0$ 的充分小的邻域 $U\subset V$，令 $\varphi=f^{-1}\psi$，则$(U,\varphi)$是 $M$ 在 $x_0$ 处的图，$\varphi(x_0)=0$，$\varphi U$ 是 $X$ 的 0-邻域. $\forall\ x\in U$，令 $\varphi(x)=te+y$，$t\in R$，$y\in Y$，则

$$\begin{aligned}\varphi\xi_\tau(x)&=\varphi\xi_\tau\varphi^{-1}(te+y)=\varphi\xi_\tau\psi^{-1}f(te+y)\\&=\varphi\xi_\tau\xi_t\psi^{-1}(y)=\varphi\xi_{\tau+t}\psi^{-1}(y)\\&=\varphi\psi^{-1}f((\tau+t)e+y)=\tau e+\varphi(x),\end{aligned}\tag{1}$$

由此得出 $e=\frac{d}{d\tau}\varphi(\xi_\tau x)|_{\tau=0}=d\varphi_x\xi(x)$. □

(1)表明 $\varphi$ 映 $\xi$ 的轨道为 $X$ 中平行于 $e$ 的直线. 因此，$\xi$ 在非奇点邻近的轨道犹如"平行直线族"；"直化引理"之名乃出于此.

**8.5.2 引理** $\Gamma$ 的 Poincaré 映射是局部 $C^r$ 同胚，且互相局部拓扑共轭.

**证** 设 $S$ 是 $\Gamma$ 在 $x_0$ 处的截面，$P$ 是 $S$ 上的 Poincaré 映射. 因 $\xi(x_0)\neq 0$，由 8.5.1 有分解 $X=Re\oplus Y(e\neq 0)$与 $C^r$ 同胚 $\varphi:U\to I\times B$，$U$ 是 $x_0$ 的邻域，$I=(-\delta,\delta)e$，$B$ 是 $Y$ 的 0-邻域，使得 $\varphi(x_0)=0$，$\forall\ x\in U$：$d\varphi_x\xi(x)=e$. 于是$S_0\triangleq\varphi^{-1}(\{0\}\times B)$是 $\Gamma$ 在 $x_0$ 处的截面，$\pi\triangleq\varphi^{-1}q\varphi\in C^r$ 是沿轨道到 $S_0$ 上的"投影"，此处 $q:X\to Y$ 是投影. 设 $P_0$ 是 $S_0$ 上的 Poincaré 映射，其定义域为 $V_0$，则 $P_0x=\pi\xi_p(x)(x\in V_0)$，因此 $P_0=\pi\xi_p|V_0\in C^r$，$P_0^{-1}=\pi\xi_{-p}|P_0V_0$，可见 $P_0$ 在 $x_0$ 处为局部 $C^r$ 同胚. 可设 $V_0$ 适当小，使 $V=S\cap\pi^{-1}V_0$ 是 $x_0$ 在 $S$ 上的邻域且 $P$ 在 $V$ 上定

义，$\pi: V \to V_0$ 是 $C^r$ 同胚. 由 $P|V=(\pi|V)^{-1}P_0\pi|V$ 看出 $P$ 在 $x_0$ 处为局部 $C^r$ 同胚，且 $P$ 与 $P_0$ 局部拓扑共轭(8.1.5).

设 $S_1$ 是 $\Gamma$ 在 $x_1$ 处的截面，$P_1$ 是 $S_1$ 上的 Poincaré 映射. 不妨设 $x_1$ 充分接近 $x_0$，且 $S_1 \subset U$（否则令 $x_1$ 沿 $\Gamma$ 作有限次小移动）. 取 $\tau \in [0,p]$，使 $x_1=\xi_\tau(x_0)$，则 $P_2 \triangleq \xi_{-\tau}P_1\xi_\tau$ 是 $\Gamma$ 在 $x_0$ 处的 Poincaré 映射. 因 $P_1 \sim P_2$，而前段所证表明 $P, P_2$ 分别与 $P_0$ 局部拓扑共轭，故 $P_1$ 与 $P$ 局部拓扑共轭. □

依 8.5.2 之证，$\mathrm{d}P(x_0)$ 与 $\mathrm{d}P_1(x_1)$ 必线性共轭，因此 $\sigma(\mathrm{d}P(x_0))$ 由 $\Gamma$ 唯一决定而与 $x_0$ 及 $S$ 的选取无关. 这引导出以下定义：

**8.5.3 定义**　若 $P$ 是 $\Gamma$ 在 $x_0$ 处的 Poincaré 映射，$x_0$ 是 $P$ 的双曲不动点，则称 $\Gamma$ 为双曲闭轨.

关于双曲不动点的某些结果可以推广于双曲闭轨. 相当于 8.2.10 的稳定性结果是：

**8.5.4 定理**　设 $\dim M<\infty$. 若 $\Gamma$ 是 $\xi$ 的双曲闭轨，则 $\xi$ 在 $\Gamma$ 上局部 $C^r$ 结构稳定，即对 $\xi$ 的任何 $C^r$ 近似 $\eta$，存在 $\Gamma$ 的邻域 $V, V'$ 与同胚 $h: V \to V'$，使 $h$ 映 $\xi$ 的轨道为 $\eta$ 的轨道.

以上结果本质上是 8.2.10 的推广，但 $h$ 的构作有一些技术性的困难，仅将其基本思路简述于下. 取定 $x_0 \in \Gamma$ 与 $\Gamma$ 在 $x_0$ 处的截面 $S$，设 $P$ 是 $S$ 上的 Poincaré 映射. 因 $P$ 在 $x_0$ 处局部 $C^r$ 结构稳定(8.2.10)，且 $P$ 对 $\xi$ 的依赖是 $C^r$ 的，故若 $\eta$ 是 $\xi$ 的 $C^r$ 近似，则 $\eta$ 亦有双曲闭轨 $\Gamma'$，$\Gamma'$ 亦以 $S$ 为截面，且 $\eta$ 在 $S$ 上的 Poincaré 映射 $P'$ 是 $P$ 的 $C^r$ 近似. 由 8.2.10，有 $x_0$ 在 $S$ 上的邻域 $U$ 及 $U$ 上的同胚 $h$，使 $P'h=hP$. 余下的关键性步骤是，将 $h$ 扩张为定理所需的同胚. 所需扩张本质上借助于 $\xi$ 与 $\eta$ 的流完成，但其构造细节颇需技巧，其详参看[136].

用类似的方法可建立 8.4.3 的以下类似：

**8.5.5 定理**　若 $\Gamma$ 是 $\xi$ 的双曲闭轨，则 $W^s(\Gamma)$ 与 $W^u(\Gamma)$（依 8.1.4）是 $M$ 的 $C^r$ 浸入子流形，二者沿 $\Gamma$ 横截相交.

**证** 取 $\Gamma$ 的充分小的邻域 $V$，令 $W_V^s(\Gamma)=\{x: O^+(x)\subset V\}$，$W_V^u(\Gamma)$ 仿此，则可验明

$$W^s(\Gamma)=\bigcup_{k\geqslant 0}\xi_{-k}W_V^s(\Gamma), W^u(\Gamma)=\bigcup_{k\geqslant 0}\xi_k W_V^u(\Gamma).$$

因此只需指明 $W_V^s(\Gamma)$（类似地 $W_V^u(\Gamma)$）是 $M$ 的 $C^r$ 浸入子流形（参考 8.4.2 之证）.

取 $x_0\in\Gamma$ 与 $\Gamma$ 在 $x_0$ 的截面 $S$，设 $P$ 是 $S$ 上的 Poincaré 映射. 令 $W^s(x_0)=W^s(x_0,P)$，$W^u(x_0)$ 仿此. 由 8.4.2，$W^s(x_0)$ 与 $W^u(x_0)$ 是 $S$ 的 $C^r$ 浸入子流形，且二者在 $x_0$ 横截相交，令

$$Q_s=\bigcup_{0<t<2p}\xi_t W^s(x_0), Q_u=\bigcup_{0<t<2p}\xi_t W^u(x_0),$$

则 $Q_s$ 与 $Q_u$ 是 $M$ 的 $C^r$ 浸入子流形，二者沿 $\Gamma$ 横截相交，注意到 $W_V^s(\Gamma)$ 与 $W_V^u(\Gamma)$ 分别为 $\Gamma$ 在 $Q_s$ 与 $Q_u$ 中的开邻域，所需结论由之得出. □

本节的基本思想在于，在每点 $x_0\in\Gamma$，$\xi$ 沿横截面的局部行为由 Poincaré 映射描述；而沿轨道方向的变化归结为若干步相继的"平行移动".

## §6 1维半动力系统

以下三节讨论一些特殊的动力或半动力系统，这些系统以其在理论上富有启发性，或以其独特的应用价值，显示出特殊的重要性.

首先考虑应当是最简单的 1 维系统. 给定 $f\in C(\mathrm{R})$，$f$ 生成半动力系统 $\{f^n: n\geqslant 0\}$. 自 60 年代以来的一些出人意外的发现表明，即使初看起来如此简单的系统亦可能呈现出惊人的复杂性. 第一个惊人的结果属于 Sarkovskii，他的结果基于自然数的一种颇为奇特的序：

$$3\prec 5\prec 7\prec\cdots\prec 2\times 3\prec 2\times 5\prec 2\times 7\prec\cdots$$
$$\cdots\prec 2^2\times 3\prec 2^2\times 5\prec\cdots\prec 2^3\prec 2^2\prec 2\prec 1.$$

**8.6.1 Sarkovskii 定理**(1964) 设 $\mathrm{PP}(f)$ 记 $f$ 的周期点的周期

之集. 若 $m\in \mathrm{PP}(f)$, $m\lhd n$, 则 $n\in \mathrm{PP}(f)$.

要说明 8.6.1 的异乎寻常的深刻性, 只需指出它的以下惊人推论就够了.

**8.6.2 推论** 若 $3\in \mathrm{PP}(f)$, 则 $\mathbb{N}\subset \mathrm{PP}(f)$.

这一特殊结论尽管已蕴涵于 Sarkovskii 定理中, 但许久未引起注意, 直至 Li 与 Yorke[114]于 1975 年重新发现. 文献[114]因其首次提出浑沌概念而负盛名.

8.6.1 的证明完全是初等的, 但颇富技巧性(参考[209]). 下面仅给出 8.6.2 的证明.

**8.6.2 之证** 设 $\Gamma=\{x_0,x_1,x_2\}$ 是 $f$ 的 3-周期轨道, 其中 $x_1=f(x_0)$, $x_2=f(x_1)$. 适当调整编号, 总可设 $x_0<x_1<x_2$ 或 $x_0>x_1>x_2$, 下面假定前者. 令 $K=[x_0,x_1]$, $L=[x_1,x_2]$, 显然 $L\subset fK$, 形式地记成 $L\vdash K$. 类似地有 $K\vdash L\vdash L$. 令

$$a=\sup(K\cap f^{-1}(x_1)),b=\inf(K\cap f^{-1}(x_2)),$$

则 $I_1=[a,b]$ 是 $K$ 中满足 $f(I_1)=L$ 的最小子区间. 因 $I_1\vdash L$, 类似地有最小区间 $I_2\subset L$, 使 $f(I_2)=I_1$. 进而有最小区间 $I_3\subset L$: $f(I_3)=I_2$, $\cdots$, 如此得出最小区间 $I_n\subset L$: $f(I_n)=I_{n-1}(n\geqslant 2)$. 于是

$$I_n\subset L=f(I_1)=f^2(I_2)=\cdots=f^n(I_n). \tag{1}$$

由(1)易推出存在 $x\in I_n\cap \mathrm{Fix}f^n$, 使 $f^i(x)\in I_{n-i}\subset L(0\leqslant i\leqslant n-2)$, $f^{n-1}(x)\in I_1\subset K$. 下面证 $x$ 是 $f$ 的 $n$-周期点, $n\geqslant 2$, $n\neq 3$(注意由 $L\vdash L$ 直接推出 $L\cap \mathrm{Fix}f\neq\varnothing$, 因此 $1\in \mathrm{PP}(f)$ 不必证, 而 $3\in \mathrm{PP}(f)$ 为已知); 为此只需证 $f^i(x)(0\leqslant i\leqslant n-1)$ 互不相同, 而这归于证 $O(x)\cap\Gamma=\varnothing$(如此则 $f^i(x)\in(x_1,x_2)(0\leqslant i\leqslant n-2)$, $f^{n-1}(x)\in(x_0,x_1)$, $x$ 的周期必为 $n$). 若 3 不整除 $n$, 则 $\Gamma\cap \mathrm{Fix}f^n=\varnothing$, 这推出 $O(x)\cap\Gamma=\varnothing$. 若 3 整除 $n$, 则 $n\geqslant 6$, 亦得 $O(x)\cap\Gamma=\varnothing$(否则必 $x\in\Gamma$, 由此易导出矛盾). □

下面通过实例来展示 1 维系统可能有的复杂性. 看来很简单的函数

$$f_\mu(x) = \mu x(1-x), x \in \mathrm{R} \tag{2}$$

($\mu>0$ 为参数)已足以说明问题,尽管从形式上看 $f_\mu$ 并不是"高度非线性的". 下面的讨论要反复用到以下简单命题:

**8.6.3 引理** 设 $f\in C^1(\mathrm{R})$, $f(x_0)=x_0$, $I_0=(a,x_0)$, $I_1=(x_0,b)$, $a<x_0<b$. 若 $f(I_0)\subset I_0$, $\forall\ x\in I_0$: $f'(x)<1$[$f'(x)>1$],则 $\forall\ x\in I_0$: $f^n(x)\uparrow x_0$[$f^n(x)\downarrow a$]. 若 $f(I_1)\subset I_1$, $\forall\ x\in I_1$: $f'(x)<1$[$f'(x)>1$],则 $\forall\ x\in I_1$: $f^n(x)\downarrow x_0$[$f^n(x)\uparrow b$].

$\forall\ \mu>0$, $f_\mu$ 有不动点 0 与 $x_\mu=1-\mu^{-1}$(当 $\mu=1$ 时 $x_\mu=0$). 因 $f'_\mu(0)=\mu$, $f'_\mu(x_\mu)=2-\mu$, 故 $0<\mu<1$ 时 0 与 $x_\mu$ 分别为 $f_\mu$ 的吸点与斥点;用 8.6.3 得 $W^s(0)$($=W^s(0,f_\mu)$,下同)$=(x_\mu,\mu^{-1})$, $\forall\ x\in\mathrm{R}\setminus[x_\mu,\mu^{-1}]$: $f_\mu^n(x)\to-\infty(n\to\infty)$; $\mu>1$ 时 0 是 $f_\mu$ 的斥点; $1<\mu<3$ 时 $x_\mu$ 是 $f_\mu$ 的吸点, $\mu>3$ 时 $x_\mu$ 是斥点. 若 $\mu\geqslant1$,则 $f((1,\infty))=f((-\infty,0))=(-\infty,0)$,且在 $(-\infty,0)$ 内 $f'_\mu(x)>1$,故 $\forall\ x\in\mathrm{R}\setminus J$: $f_\mu^n(x)\to-\infty(n\to\infty)$. 可见只需考虑 $f_\mu|J$. 用 8.6.3 易得出当 $\mu=1$ 时 $W^s(0)=J$;当 $1<\mu\leqslant3$ 时 $W^s(x_\mu)=(0,1)$.

真正复杂的情况出现在 $\mu$ 越过 3 以后. 对于 $3<\mu<4$ 有以下结论:令 $\mu_1=3$;存在 $\mu_2>\mu_1$,当 $\mu_1<\mu<\mu_2$ 时 $f_\mu^2$ 有 4 个不动点,其中两个是 $f_\mu$ 的 2-周期点;又存在 $\mu_3>\mu_2$,当 $\mu_2<\mu<\mu_3$ 时 $f_\mu^4$ 有 8 个不动点,其中 4 个是 $f_\mu$ 的 4-周期点. 如此得到序列 $\{\mu_n\}$, $\mu_n\to\mu_\infty=3.569945672\cdots$;当 $\mu$ 接近 $\mu_\infty$ 时 $f_\mu$ 的周期点无限增多.

当 $\mu$ 越过 4 之后有一些更引人的结论.

**8.6.4 定理** 若 $\mu>4$,则存在 $J$ 的完备稀疏子集 $\Lambda$(称为 Cantor 集),使得 $f_\mu(\Lambda)\subset\Lambda$; $\forall\ x\in\mathrm{R}\setminus\Lambda$: $f_\mu^n(x)\to-\infty(n\to\infty)$.

**证** 固定 $\mu>4$,令 $f=f_\mu$. 因 $f(1/2)=\mu/4>1$,故有一中心在 1/2 的开区间 $A_0$: $f(A_0)>1$; $J\setminus A_0$ 由互不相交的闭区间 $I_0$, $I_1$ 组成, $f(I_i)=J(i=0,1)$. 令 $A_1=J\cap f^{-1}(A_0)$,则 $A_1$ 由两个分别含于 $I_0$, $I_1$ 的开区间组成. 归纳地,

$$A_n \triangleq J\cap(f^n)^{-1}(A_0)\setminus(A_0\cup A_1\cdots\cup A_{n-1})$$

由 $2^n$ 个互不相交的开区间组成. 令 $\Lambda=J\setminus\bigcup_0^\infty A_n$，则显然 $\Lambda$ 是完备集，$f(\Lambda)\subset\Lambda$；$\forall\ x\in\mathrm{R}\setminus\Lambda$：$f^n(x)\to-\infty(n\to\infty)$. 余下只需证 $\Lambda$ 为疏集. 只就 $\mu>2+\sqrt{5}$ 的情况证明；当 $4<\mu\leqslant2+\sqrt{5}$ 时证明更精细些. 直接计算可验知在 $J\setminus A_0$ 上 $|f'(x)|>1$，因此 $\exists\ \lambda>1$，在 $\Lambda$ 上 $|f'(x)|>\lambda$，从而 $|f^{n\prime}(x)|>\lambda^n$. 若 $\Lambda$ 非疏集，则有 $[x,y]\subset\Lambda, x<y$，于是由中值定理推出对大的 $n$ 有 $|f^n(x)-f^n(y)|\geqslant\lambda^n|x-y|>1$，得出矛盾. □

考察 8.6.4 之证不难看出，$f_\mu^n$ 的图形与直线 $y=x$ 至少交 $2^n$ 次，这推出 $f_\mu$ 至少有 $2^n$ 个 $n$-周期点. 从 $A_n$ 及 $\Lambda$ 的构成看出 $\overline{\mathrm{Per}f_\mu}=\Lambda(\mu>4)$.

**8.6.5 定义** 设 $M$ 为度量空间，$f\in C(M,M)$. 若对任何非空开集 $U,V\subset M$，$\exists\ n>0$：$V\cap f^nU\neq\varnothing$，则说 $f$ 是拓扑可迁的. 若有 $\varepsilon>0$，对任给 $x\in M$ 与 $x$ 的邻域 $V$，$\exists\ y\in V, n\geqslant1$：$\mathrm{d}(f^n(x),f^n(y))>\varepsilon$，则说 $f$ 敏感依赖于初始条件. 若 $f$ 拓扑可迁、敏感依赖于初始条件且 $\overline{\mathrm{Per}f}=M$，则说 $f$ 是浑沌的.

**注** 若 $f$ 有一稠密轨道，则 $f$ 必拓扑可迁.

**8.6.6 定理** 设 $f_\mu$ 依(2). 若 $\mu>2+\sqrt{5}$，则 $f_\mu$ 在 $\Lambda$(依 8.6.4)上是浑沌的.

**证** 已指明 $\overline{\mathrm{Per}f_\mu}=\Lambda$. 设 $0<\varepsilon<\mathrm{mes}A_0$(记号依 8.6.4 之证). 若 $x,y\in\Lambda, x<y$，则在 $(x,y)$ 中必有 $\Lambda$ 的余区间，这意味着 $\exists\ n\geqslant1$：$f_\mu^n(x)$ 与 $f_\mu^n(y)$ 在 $A_0$ 的两侧，因此 $|f_\mu^n(x)-f_\mu^n(y)|>\varepsilon$. 这表明 $f_\mu$ 敏感依赖于初始条件. 其次，可指明 $f$ 在 $\Lambda$ 中有一稠密轨道(这一事实的证明是初等的，但颇琐碎，从略)，因此 $f_\mu$ 在 $\Lambda$ 上是浑沌的. □

## §7 平面动力系统

本节考虑 $\mathrm{R}^2$ 中的动力系统. 这类系统具有较好的性质主要有赖于以下著名拓扑定理：

**8.7.1 Jordan 曲线定理** $R^2$ 中任何闭 Jordan 曲线(即 $S^1$ 的同胚象)的补集是互不相交的两区域,即曲线的内部与外部.

**注** Jordan 曲线定理在球面 $S^2$ 上亦成立. 实际上,本节结果亦适用于球面动力系统.

以下给定 $C^1$ 向量场 $\xi:\Omega\subset R^2\to R^2$. 任给 $x\in\Omega$,令 $x(t)=\xi_t(x)$, $O(x)=O(x,\xi)$;$O^+(x)$,$\omega(x)$等仿此. 令 $E=\{x:\xi(x)=0\}$. 称线段 $L$ 为 $\xi$ 的截线,若$\forall\ x\in L$:$L$ 与 $\xi(x)$不共线. 在本节中,截线是一个基本的辅助工具,有关它的性质汇集于下:

**8.7.2 命题** 设 $L$ 是 $\xi$ 的截线,$x\in\Omega$. (i)$L\cap E=\varnothing$;(ii)若 $O(x)\cap L\neq\varnothing$,则 $O(x)$从同一侧穿过 $L$;(iii)单调性:若 $x_i=x(t_i)\in L$,$i=1,2,3$,$t_1<t_2<t_3$,则 $x_2\in[x_1,x_3]$;(iv)$\omega(x)\cap L$ 至多含一点.

**证** (i)(ii)是明显的.

(iii)不妨设$\forall\ t\in(t_1,t_2)$:$x(t)\overline{\in}L$,于是自 $x_1$ 至 $x_2$ 的一段轨道与线段$[x_1,x_2]$一起组成一闭 Jordan 曲线 $\Gamma$. 设在点 $x_2$ 处 $\xi(x_2)$指向 $\Gamma$ 的外[内]部. 因 $\Gamma$ 外[内]部的轨道不能进入 $\Gamma$ 的内[外]部,故必 $x_2\in[x_1,x_3]$.

(iv)设 $\omega(x)\cap L$ 含两点 $y$,$z$,取 $y$,$z$ 的互不相交小邻域 $U$,$V$. 因 $O^+(x)$无限次进入 $U$,$V$,而据 8.5.1 易见进入 $U$,$V$ 的轨道必穿过 $L$,这就必然会与已证的单调性结论冲突. □

以下是关于平面动力系统的基本定理.

**8.7.3 Poincaré-Bendixson 定理** 设 $x\in\Omega$,$\overline{O^+(x)}$是 $\Omega$ 的紧子集. (i)若 $E\cap\omega(x)=\varnothing$,则 $\omega(x)$是一闭轨. (ii)若 $\omega(x)=\{y\}$,则 $x(t)\to y\in E(t\to\infty)$. (iii)若 $A\triangleq E\cap\omega(x)=\{x_1,x_2,\cdots,x_n\}$,$n\geqslant1$,$\omega(x)$非单点集,则 $\omega(x)\backslash A$ 由至多可数条轨道组成,其中每条两端分别趋近于 $A$ 中某点.

**证** (i)任取 $y\in\omega(x)$,今指明 $O(y)$为闭轨且 $\omega(x)=O(y)$. 因 $\omega(x)$是紧不变集,故$\overline{O(y)}\subset\omega(x)$,从而$\varnothing\neq\omega(y)\subset\omega(x)$. 任取 $z\in\omega(y)$,必 $z\overline{\in}E$,于是由 8.5.1 知 $\xi$ 有过 $z$ 的截线 $L$. 由 8.7.2(iv)有 $\omega$

$(x)\cap L=\{z\}$，从而 $O(y)\cap L\subset\{z\}$. 因 $O^+(y)$ 无限次到达 $z$ 邻近，从而无限次经过 $z$，故 $O(y)$ 为闭轨，设其周期为 $p$. 取 $t_n\to\infty$，使 $x_n=x(t_n)\in L$，$x_n\to z$，当 $t_n<t<t_{n+1}$ 时 $x(t)\bar{\in} L$. 由流的连续性推出 $t_{n+1}-t_n\to p$. 设在 $\overline{O^+(x)}$ 上及其邻近 $|\xi'(u)|\leqslant C$，据此用一标准的 Gronwall 不等式论证可得出

$$|x_n(t)-z(t)|\leqslant e^{2pC}|x_n-z|(0\leqslant t\leqslant 2p),$$

这与 $x_n\to z, t_{n+1}-t_n\to p$ 一起推出 $\omega(x)\subset O(z)$，因此 $\omega(x)=O(z)=O(y)$，$\omega(x)$ 是闭轨.

(ii)是明显的.

(iii)因 $\omega(x)$ 连通(8.1.2)，且多于一点，故 $\omega(x)\neq A$. 任取 $y\in\omega(x)\backslash A$. 若有 $z\in\omega(y)\backslash E$，则用(i)的证法可推出 $\omega(x)=O(z)$，与 $\omega(x)\cap E\neq\varnothing$ 矛盾. 因此 $\omega(y)\subset E\cap\omega(x)=A$. 设 $x_i\in\omega(y)$，则必 $\omega(y)=\{x_i\}$，否则对 $x_i$ 的某个小邻域 $U$，$O^+(y)$ 无限次穿过紧集 $\partial U$，这得出 $\omega(y)\cap\partial U\neq\varnothing$，与 $\omega(y)\subset A$ 矛盾. 于是由(ii)有 $y(t)\to x_i(t\to\infty)$. 同理 $y(t)\to x_j\in A(t\to-\infty)$. 若 $x_i=x_j$，则 $\Gamma\triangleq O(y)\cup\{x_i\}$ 是一闭 Jordan 曲线，这样的 $\Gamma$ 不能内部相交，因此至多可数个. 若 $x_i\neq x_j$，则 $O(y)$ 是两端分别趋于 $x_i$ 与 $x_j$ 的唯一轨道. 否则有轨道 $O(u)\subset\omega(x)$：$u(t)\to x_i, u(-t)\to x_j(t\to\infty)$，$C\triangleq O(y)\cup O(u)\cup\{x_i,x_j\}$ 是一闭 Jordan 曲线. 取 $\xi$ 的截线 $L$、$L_1$，使 $y\in L, u\in L_1$. 设 $O^+(x)$ 先交 $L$ 于点 $a$，后交 $L_1$ 于点 $b$，不妨设 $a,b$ 皆在 $C$ 之内部. 自 $a$ 到 $b$ 的一段轨道与线段 $[y,a]$，$[b,u]$ 及 $O^+(u)$，$O^+(y)$，$x_i$ 一起构成闭 Jordan 曲线 $K$，而 $O^+(x)$ 最终进入 $K$ 之内部，这与 $x_j\in\omega(x)$ 矛盾. 综上所证得出(iii). □

在 $\omega(x)$ 仅含有限多个奇点的情况下，8.7.3 概括了所有的可能性. 不过，$\omega(x)$ 亦可能含无限多个奇点，这种情况不拟讨论.

设 $\Gamma$ 是 $\xi$ 的闭轨. 可用§5 中的方法研究 $\Gamma$. 不过，$\Gamma$ 是平面闭 Jordan 曲线这一事实引导出一些特殊的概念与方法. 任给 $\epsilon>0$，令

$$N_\varepsilon(\Gamma) = \{x: x \text{ 在 } \Gamma \text{ 之外部且 } d(x,\Gamma) < \varepsilon\}.$$

若$\forall\ \varepsilon>0$，$\exists\ \delta>0$，$\forall\ x\in N_\delta(\Gamma)$：$O^+(x)\subset N_\varepsilon(\Gamma)$[且 $\omega(x)=\Gamma$]，则说 $\Gamma$ 外侧轨道稳定[外侧渐近轨道稳定]. 类似地可定义内侧[渐近]轨道稳定.

**8.7.4 引理** 若 $\Gamma=O(x)$是 $\xi$ 的闭轨，其内部 $G$ 含于 $\Omega$，则 $\xi$ 在 $G$ 内有奇点.

**证** 设 $\xi=(u,v)$，$x(t)=(x_1(t),x_2(t))$，则 $x'_1(t)=u(x(t))$，$x'_2(t)=v(x(t))$. 设 $p$ 是 $\Gamma$ 的周期. 若 $\xi$ 在 $G$ 内无奇点，则

$$0 = \int_\Gamma \frac{u\mathrm{d}v - v\mathrm{d}u}{u^2+v^2}$$

$$= \int_0^p \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{arctg}\,\frac{x'_2(t)}{x'_1(t)}\mathrm{d}t = \pm 2\pi,$$

得出矛盾. □

**8.7.5 定理** 对 $\xi$ 的闭轨 $\Gamma=O(x)$有以下结论：(i) $\Gamma$ 外侧渐近轨道稳定$\Leftrightarrow$存在 $\Gamma$ 外部的点 $y$ 使 $\Gamma=\omega(y)$. (ii)若 $\Gamma$ 外侧轨道稳定而非渐近稳定，则$\forall\ \varepsilon>0$，$N_\varepsilon(\Gamma)$内含有 $\xi$ 的异于 $\Gamma$ 的闭轨.

**证** (i)设 $\Gamma=\omega(y)$，$y$ 在 $\Gamma$ 之外部. 取 $\xi$ 的过 $x$ 的截线 $L$. 由 $x\in\omega(y)$有 $t_n\to\infty$：$y_n=y(t_n)\in L$，$y_n\to x$；可设$\forall\ t\in(t_n,t_{n+1})$：$y(t)\bar{\in} L$. 从 $y_n$ 到 $y_{n+1}$的一段轨道与$[y_n,y_{n+1}]$一起组成闭 Jordan 曲线 $\Gamma_n$，以 $G_n$ 记 $\Gamma_n$ 之内部，令 $D_n=G_n\setminus\overline{G}_{n+1}$. 给定 $\varepsilon>0$，取 $n$ 充分大，使 $D_n\subset N_\varepsilon(\Gamma)$且 $E\cap\bigcup_n^\infty D_k=\varnothing$. 取 $\delta>0$ 充分小，设 $z\in N_\delta(\Gamma)$，可设 $z\in D_n$，$z\bar{\in}O(y)$. $O^+(z)$必穿过$[y_{n+1},y_{n+2}]$进入 $D_{n+1}$，否则$\overline{O^+(z)}\subset\overline{D_n}$，由 8.7.3(i)推出 $\omega(x)$为闭轨，且 $\omega(z)\subset D_n$；由 8.7.4 推出 $\omega(z)$内部有 $\xi$ 的奇点，这与 $E\cap D_n=\varnothing$矛盾. 因此 $O^+(z)$将相继进入每个 $D_k$ $(k>n)$，从而 $\omega(z)=\Gamma$. 这表明 $\Gamma$ 外侧渐近轨道稳定. 逆命题不必证。

(ii)可设 $E\cap\overline{N_\varepsilon(\Gamma)}=\varnothing$. 取 $y$，使 $O^+(y)\subset N_{\varepsilon/2}(\Gamma)$，则 $E\cap\omega(y)=\varnothing$. 由 8.7.3(i)，$\omega(y)$为闭轨. 由已证的(i)，$\omega(y)\neq\Gamma$，于是 $\omega$

$(y)$是 $N_\epsilon(\Gamma)$内与 $\Gamma$ 相异的闭轨. □

**注** 对于 $\Gamma$ 的内侧轨道稳定性,显然可建立与 8.7.5 相对应的结果.

若闭轨 $\Gamma$ 的某邻域内无其它闭轨,则称 $\Gamma$ 为极限环;称内外两侧皆轨道稳定的极限环为稳定极限环. 由 8.7.5,稳定极限环必(双侧)渐近轨道稳定. 以 $-\xi$ 替代 $\xi$,可得出不稳定概念及相应的结论.

现在结合运用§5 中的方法. 设 $\Gamma=O(x_0)$是 $\xi$ 的闭轨. 取 $\xi$ 过 $x_0$ 的截线 $L$,在 $L$ 上导入坐标 $s$,使 $x_0$ 对应 $s=0$. 设 $P$ 是 $L$ 上的 Poincaré 映射,则 $s=0$ 是 $P$ 的不动点,$P'(0)\geqslant 0$. 据 8.5.3,当 $P'(0)\neq 0,1$ 时 $\Gamma$ 为双曲闭轨. 其次,显然 $\Gamma$ 是极限环$\Leftrightarrow$0 是 $P$ 的孤立不动点. 下面给出用 $P$ 的导数判定 $\Gamma$ 为稳定或不稳定极限环的方法.

**8.7.6 定理** 设 $\xi\in C^r$,$r$ 足够大. (i)若 $P'(0)<1[P'(0)>1]$,则 $\Gamma$ 是稳定[不稳定]极限环. (ii)若 $P'(0)=1,P^{(i)}(0)=0(1<i<k)$,$P^{(k)}(0)\neq 0,k>1$,则当 $k$ 为奇数,$P^{(k)}(0)<0[P^{(k)}(0)>0]$时 $\Gamma$ 是稳定[不稳定]极限环,当 $k$ 为偶数时 $\Gamma$ 是半稳定极限环(半稳定意味着一侧稳定,另一侧不稳定).

**证** (i)直接由 8.5.5 推出.

(ii) 将 $P(s)$在 $s=0$ 处展开:

$$P(s)=s+\frac{1}{k!}P^{(k)}(0)s^k+0(s^k).$$

由此可见 $s=0$ 是孤立不动点,从而 $\Gamma$ 是极限环. 若 $k$ 为奇数,则 $\operatorname{sgn}(P(s)-s)=\operatorname{sgn}P^{(k)}(0)\operatorname{sgn}s$($|s|$充分小);于是当 $P^{(k)}(0)<0[P^{(k)}(0)>0]$时 $\Gamma$ 稳定[不稳定]. 若 $k$ 为偶数,则对小的 $s$ 有 $\operatorname{sgn}(P(s)-s)=\operatorname{sgn}P^{(k)}(0)$,于是 $\Gamma$ 半稳定. □

## §8 单调系统

近十多年来,具有明显生态与经济学背景的"单调动力系统"受

到广泛注意.主要由于 Hirsch,Smith 等人的工作(如[72～74,169～173]),在这一方向出现了一系列深刻结果,本节意在对其基本部分加以初步概括.

设 $X$ 中已由序锥 $X_+$ 导入序 $\leqslant$,记号 $<$ 与 $\ll$ 依 § 2.1. 给定 $C^1$ 向量场 $\xi:\Omega\subset X\to Y$,记 $E=\{x:\xi(x)=0\}$,$x(t)=\xi_t(x)$,$O(x)=O(x,\xi)$,$O^+(x)$,$\omega(x)$ 等仿此.

**8.8.1 定义** 若当 $x<y$,$t>0$,$x(t)$,$y(t)$ 有定义时 $x(t)\leqslant y(t)$ [$x(t)\ll y(t)$],则称 $\xi$ 为单调[强单调]系统,或称 $\xi_t$ 为单调[强单调]流.

以下是单调系统的两个重要例子.

**8.8.2 例** 设 $\xi=(\xi_i)\in C^1(\mathrm{R}^n,\mathrm{R}^n)$. 今指明 $\xi$ 单调 $\Leftrightarrow\forall\ x\in\mathrm{R}^n$: $\xi'(x)$ 本性非负(即其非对角线元非负). 首先,若 $\xi$ 单调,取定 $x\in\mathrm{R}^n$,令 $M(t)=\xi'_t(x)$,则

$$M(t)h=\frac{\mathrm{d}}{\mathrm{d}s}\xi_t(x+sh)\big|_{s=0}\geqslant 0\ (t>0,h\in\mathrm{R}^n_+),$$

可见 $M(t)\geqslant 0(t>0)$. 而 $M(0)=I$,$\xi'(x)=M'(0)$(参考 8.2.1),故知 $\xi'(x)$ 本性非负. 其次,设 $\forall\ x\in\mathrm{R}^n$:$\xi'(x)$ 本性非负,$x,y\in\mathrm{R}^n$,$x\leqslant y$. 取充分小的 $\varepsilon>0$,令 $\xi^\varepsilon=\xi+\varepsilon e$,$e=(1,1,\cdots,1)\in\mathrm{R}^n$;以 $y^\varepsilon(t)$ 记 IVP

$$u'=\xi^\varepsilon(u),u(0)=y+\varepsilon e$$

的解. 注意 $x(0)=x\ll y+\varepsilon e=y^\varepsilon(0)$. 对任何使 $x(t)$,$y^\varepsilon(t)$ 有定义的 $t\geqslant 0$,必有 $x(t)\ll y^\varepsilon(t)$,否则有 $\tau>0$,$\forall\ t\in[0,\tau)$:$x(t)\ll y^\varepsilon(t)$,对某个 $i$ 有 $x_i(\tau)=y^\varepsilon_i(\tau)$,从而 $x'_i(\tau)\geqslant y^{\varepsilon\prime}_i(\tau)$. 另一方面,记 $u=y^\varepsilon(\tau)-x(\tau)$,则

$$\begin{aligned}y^{\varepsilon\prime}_i(\tau)&=\xi_i(y^\varepsilon(\tau))+\varepsilon>\xi_i(y^\varepsilon(\tau))\\&=\xi_i(x(\tau))+\int_0^1\xi'_i(x(\tau)+su)u\mathrm{d}s\\&\geqslant\xi_i(x(\tau))=x'_i(\tau),\end{aligned}$$

得出矛盾. 因此 $x(t)\ll y^{\epsilon}(t)$；令 $\epsilon\downarrow 0$ 得 $x(t)\leqslant y(t)(t\geqslant 0)$，可见 $\xi$ 是单调的.

更精细的分析可以证明：若 $\forall\, x\in\mathbf{R}^n$：$\xi'(x)$ 本性非负且不可约(2.3.1)，则 $\xi$ 是强单调的(见[170]).

**8.8.3 例** 设 $X=C([-r,0],\mathbf{R}^n)$，$X_+=C([-r,0],\mathbf{R}^n_+)$，$r>0$ 是给定的. 设 $f\in C^1(X,\mathbf{R}^n)$ 满足条件

(C) $\forall\, \varphi\in X, u\in X_+$，当 $u_i(0)=0$ 时 $f'_i(\varphi)u\geqslant 0$. 以 $\xi_t(\varphi)$ 记如下泛函微分方程 IVP 的解(参考[169])：

$$x'(t)=f(x_t),x_0=\varphi\in X,$$

今指明 $\xi_t$ 是 $X$ 上的单调流. 设 $\varphi,\psi\in X$，$\varphi\leqslant\psi$，$x(t)=\xi_t(\varphi)$ 与 $y(t)=\xi_t(\psi)$ 在 $[-r,\rho)$ 上定义. 令 $f^{\epsilon}=f+\epsilon e$，$\psi^{\epsilon}=\psi+\epsilon e$，$e$ 如 8.8.2；以 $y^{\epsilon}(t)$ 记 IVP"$z'(t)=f^{\epsilon}(z_t),z_0=\psi^{\epsilon}$"的解. 任给 $\sigma\in(0,\rho)$，取充分小的 $\epsilon>0$，使 $y^{\epsilon}(t)$ 在 $[-r,\sigma]$ 上定义. 若有 $\tau\in(0,\sigma]$，使得 $\forall\, t\in[0,\tau)$：$x(t)\ll y^{\epsilon}(t)$，对某个 $i$ 有 $x_i(\tau)=y^{\epsilon}_i(\tau)$，则 $x'_i(\tau)\geqslant y^{\epsilon\prime}_i(\tau)$. 另一方面，$u\triangleq(y)^{\epsilon}-x)_{\tau}\geqslant 0$，$u_i(0)=0$，于是由条件(C)有

$$\begin{aligned} y^{\epsilon\prime}_i(\tau) &= f_i(y^{\epsilon}_{\tau})+\epsilon>f_i(y^{\epsilon}_{\tau}) \\ &= f_i(x_{\tau})+\int_0^1 f'_i(x_{\tau}+su)u\,\mathrm{d}s\geqslant x'_i(\tau), \end{aligned}$$

得出矛盾. 可见 $\forall\, t\in[0,\sigma]$：$x(t)\ll y^{\epsilon}(t)$. 令 $\epsilon\downarrow 0$，然后令 $\sigma\to\rho$，得出 $x(t)\leqslant y(t)(\forall\, t\in[0,\rho])$.

**8.8.4 定理**(Selgrade, 1980) 设 $\xi$ 单调，$x\in\Omega$. 若 $\xi(x)0\geqslant[\xi(x)\leqslant 0]$，则 $x(t)$ 在 $t\geqslant 0$ 时单调增[减].

**证** 设 $\xi(x)\geqslant 0$，$t\geqslant 0$，则

$$\begin{aligned} x'(t) &= \frac{\mathrm{d}}{\mathrm{d}s}x(t+s)\big|_{s=0}=\frac{\mathrm{d}}{\mathrm{d}s}\xi_t(x(s))\big|_{s=0} \\ &= \xi'_t(x)x'(0)=\xi'_t(x)\xi(x) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\xi_t(x+s\xi(x))\big|_{s=0}\geqslant 0, \end{aligned}$$

这得出 $x(t)$ 单调增. $\xi(x)\leqslant 0$ 的情况可类似处理. □

以上结果表明,单调系统的解在初始时刻增加就一直保持增加. 这一事实被称为"点火原理",它是判定收敛性的重要依据. 由 8.8.4 直接得出:

**8.8.5 推论** 设 $\xi$ 单调. 若 $x\leqslant y,\xi(y)\leqslant 0$,则 $O^+(x)\leqslant y$, $\omega(x)\leqslant y$. 若 $x\geqslant z,\xi(z)\geqslant 0$,则 $O^+(x)\geqslant z$, $\omega(x)\geqslant z$. 若 $z\leqslant x\leqslant y$, $\xi(z)\geqslant 0,\xi(y)\leqslant 0$,则 $z\leqslant\omega(x)\leqslant y$.

**8.8.6 引理** 设 $\xi$ 单调, $\overline{O^+(x)}$ 是 $\Omega$ 的紧子集. (i)若 $\omega(x)$ 非闭轨,则 $O^+(x)$ 中没有依 $<$ 相关的点. (ii) $\omega(x)$ 中没有依 $\ll$ 相关的点.

**证** (i)设有 $0\leqslant s<\tau$,使 $x(s)<x(\tau)$,不妨设 $s=0$(否则以 $x(s)$ 代 $x$). 由单调性推出 $x(k\tau)\leqslant x((k+1)\tau)$ $(k\geqslant 0)$,于是 $x(k\tau)\to y(k\to\infty)$,且 $y=y(\tau)$. 今证 $\omega(x)=O(y)$. 显然 $O(y)\subset\omega(x)$. $\forall z\in\omega(x)$,设 $x(t_i)\to z(t_i\to\infty)$. 可设 $t_i=k_i\tau+s_i,k_i\in Z_+,s_i\to s\in[o,\tau]$,于是 $x(t_i)=\xi_{s_i}(x(k_i\tau))\to\xi_s(y)=y(s)$,可见 $z=y(s)\in O(y)$.

(ii)若 $\omega(x)$ 中有依 $\ll$ 相关的点,则 $O^+(x)$ 中亦有依 $\ll$ 相关的点,不妨设 $x\ll x(\tau),\tau>0$. 由已证之(i),有 $\omega(x)=O(y),y=y(\tau)$. 因 $O(y)$ 亦必以任何邻近 $\tau$ 的 $\tau'$ 为周期,故必 $O(y)=\{y\}$,得出矛盾. □

**8.8.7 择一定理**(Hirsch, 1985) 设 $\xi$ 强单调, $x<y$, $\overline{O^+(x)}$ 与 $\overline{O^+(y)}$ 是 $\Omega$ 的紧子集,则以下命题两择一:(P) $\omega(x)\ll\omega(y)$; (Q) $\omega(x)=\omega(y)\subset E$.

**证** 首先设 $\omega(x)=\omega(y)$,证 $\omega(x)\subset E$. 设 $t_k\to\infty,x(t_k)\to u$, $y(t_k)\to v$;显然 $u\leqslant v$. 若 $u<v$,则 $u(t)\ll v(t)$ $(t>0)$;但 $u(t),v(t)\in\omega(x)$,与 8.8.6(ii)矛盾. 因此 $u=v$. $\forall t>0$,有 $x(t)\ll y(t)$,于是对充分小的 $\delta>0$ 有 $x(t+\delta)\leqslant y(t)$,从而 $x(t_k+t+\delta)\leqslant y(t_k+t)$,这推出 $u(t+\delta)\leqslant u(t)$. 同理 $u(t)\leqslant u(t+\delta)$,因此 $u(t)\equiv u\in E$. 这证得 $\omega(x)\subset E$.

其次设(Q)不真,则必 $\omega(x)\neq\omega(y)$,不妨设 $\omega(x)\not\subset\omega(y)$,今证

(P)必为真. 分以下几个步骤：

1° 设 $x(t_k)\to u\in\omega(x)\backslash\omega(y)$，$y(t_k)\to v\in\omega(y)$，$t_k\to\infty$. 必定 $u<v$，不妨设 $u\ll v$（否则代以 $u(t)$，$v(t)$，$t>0$），今证 $O^+(u)\leqslant v$（从而 $\omega(u)\leqslant v$）. a 若 $O^+(u)\not\leqslant v$，则有最大的 $\tau>0$，在$[0,\tau]$上 $u(t)\leqslant v$. 因 $v\overline{\in}O^+(u)$（否则 $u\in\omega(y)$!），故$\forall\ t\in[0,\tau]$：$u(t)<v$，$u(1+t)\ll v(1)$. 于是$\exists\ \varepsilon>0$，$\forall\ t\in[0,\tau+\varepsilon]$：$u(1+t)\ll v(1)$，从而 $x(t_k+t)\leqslant y(t_k)$（$k$ 充分大），这推出 $u(t)\leqslant v(t\in[0,\tau+\varepsilon])$，得出矛盾.

2° 证 $\omega(u)\ll\omega(y)$，$u$，$v$ 依上段. 可设 $\omega(u)\ll v$（否则以 $v(1)$ 代 $v$），于是$\exists\ \tau>0$：$\omega(u)\ll y(\tau)$；从而$\forall\ t\geqslant\tau$：$\omega(u)\ll y(t)$，这推出 $\omega(u)\leqslant\omega(y)$. 必 $\omega(u)<\omega(y)$（否则$\exists\ z\in\omega(u)\cap\omega(y)$：$z\ll v$，与 8.8.6(ii)矛盾）. 用 $\xi_t(t>0)$作用后得 $\omega(u)\ll\omega(y)$.

3° 取 $z\in\omega(u)$. 由 $z\ll\omega(y)$推出有 $\tau>0$：$x(\tau)\ll\omega(y)$，由此易推出 $\omega(x)\leqslant\omega(y)$. 必定 $\omega(x)<\omega(y)$（否则$\exists\ q\in\omega(x)\cap\omega(y)$：$z\ll q$，与 8.8.6(ii)矛盾），从而 $\omega(x)\ll\omega(y)$，即命题(P)成立. □

**8.8.8 定理**(Hirsch，1985)　设 $X=\mathbf{R}^n$，$\xi$ 强单调，$E$ 可数，$W=\{x:\overline{O^+(x)}\subset\Omega$ 紧$\}$，则对几乎所有 $x\in W$，当 $t\to\infty$时 $x(t)$收敛于 $E$ 中某点.

**证**　注意 $x(t)\to u(t\to\infty)\Leftrightarrow\omega(x)=\{u\}$. 令

$$Q=\{x\in W:\forall\ u\in E,\omega(x)\neq\{u\}\},$$

今证 $\mathrm{mes}Q=0$. 取定 $v\gg0$，任给平行于 $v$ 的直线 $L$，$L$ 依$\ll$为全序集，今证 $L\cap Q$ 可数（由此及 Fubini 定理得 $\mathrm{mes}Q=0$）. $\forall\ x\in W$，由 $\omega(x)$连通而 $E$ 可数，知 $\omega(x)\subset E\Rightarrow x\overline{\in}Q$. 因此当 $x,y\in L\cap Q$，$x\ll y$ 时必 $\omega(x)\ll\omega(y)$（用 8.8.7!）. 这样，$\{\omega(x):x\in L\cap Q\}$依$\ll$为全序集. $\forall\ x\in L\cap Q$，取 $u_x\in\omega(x)$. 若 $L\cap Q$ 不可数，则必有 $y\in L\cap Q$，$x_i\in L\cap Q$：$u_{x_i}\to u_y$，不妨设 $x_i\gg y$，这推出 $\omega(y)\leqslant u_y$，从而$\omega(y)=\{u_y\}$，这与 $y\in Q$ 矛盾. □

# 参考文献


[1]Adams,R. A.,Sobolev Spaces,Acad. Press,1975.

[2]Amann,H.,Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces,SIAM Rev.,**18**(1976),620—709.

[3]Amann,H.,Nonlinear Operators in Ordered Banach Spaces and Some Applications to Nonlinear Boundary Value Problems,New York:Springer-Verlag,1976.

[4]Ambrosetti,A.,Rabinowitz,P. H.,Dual variational methods in critical point theory and applications,J. Funct. Anal.,**14**(1973),349—381.

[5]Ang,D. D.,Schmitt,K.,Vy,L. K.,Noncoercive variational inequalities:some application,Nonlinear Anal.,**15**(1990),497—512.

[6]Arcoya,D.,Cañada,A.,Critical point theorems and applications to nonlinear boundary value problems,Nonlinear Anal.,**14**(1990),393—411.

[7]Attouch,H.,On the maximality of the sum of two maximal monotone operators,Nonlinear Anal.,**5**(1981),143—147.

[8]Attouch,H.,Moudafi,A.,Riahi,H.,Quantitative stability analysis for maximal monotone operators and semigroups of contractions,Nonlinear Anal.,**21**(1993),697—723.

[9]Auchmuty,G.,Variational principles for operator equations and initial value problems,Nonlinear Anal.,**12**(1988),531—564.

[10]Auslender,A.,Numerical methods for nondifferentiable convex optimization,Math. Prog.,**30**(1987),102—126.

[11]Avriel,M.,Nonlinear Programming,Analysis and Methods,Prentice Hall,1976.

[12]Bacciotti,A.,Kalouptsidis,N.,Topological dynamics of control systems:stability and attraction,Nonlinear Anal.,**10**(1986),547—565.

[13]Banas ,J.,El-Sayed,W. G.,Measures of noncompactness and solvability of an integral equation in the class of functions of locally bounded

variation, J. Math. Anal. Appl., 167(1992), 133—151.

[14] Barbu, V., Nonlinear Semigroups and Differential Equations in Banach Spaces, Noordhoff Int. Dubl., 1976.

[15] Bartolo, P., Benci, V., Fortunato, D., Abstract critical point theorems and applications to some nonlinear problems with "strong" resonance at infinity, Nonlinear Anal., 7(1983), 981—1012.

[16] Benavides, T. D., Set-contractions and ball-contractions in some classes of spaces, J. Math. Anal. Appl., 136(1988), 131—140.

[17] Benci, V., Some critical point theorems and applications, Comm. Pure Appl. Math., 33(1980), 147—172.

[18] Benci. V., A geometrical index for the group $S^1$ and some applications to the study of periodic soluctions of ordinary differential equations, Comm. Pure Appl. Math., 34(1981), 393—432.

[19] Benci, V., Pacella, F., Morse theory for symmetric functionals on the sphere and an application to a bifurcation problem, Nonlinear Anal., 9(1985), 763—773.

[20] Berger, M. S., Nonlinearity and Functional Analysis, Acad. Press, 1977.

[21] Berman, A., Plemmons, R., Nonnegative Matrices in the Mathematical Sciences, Acad. Press, 1979.

[22] Bhakta, P. C., Roychaudhuri, S., Optimization in Banach spaces, J. Math. Anal. Appl., 134(1988), 460—470.

[23] Bhatia, N. P., Egerland, W. O., On the existence of Li-Yorke points in the theory of chaos, Nonlinear Anal., 10(1986), 541—545.

[24] Bobylev, N. A., Burman, Y. M., Morse lemmas for multidimensional variational problems, Nonlinear Anal., 18(1992), 595—604.

[25] Borwein, J. M., Strojwas, H. M., The hypertangent cone, Nonlinear Anal., 13(1989), 125—144.

[26] Borwein, J. M., Wolkowicz, H., A simple constraint qualification in infinite dimensional programming, Math. Prog. Study, 35(1986), 83—96.

[27] Browder, F. E., Fixed point theory and nonlinear problems, Bull. Amer. Math. Soc., 9(1983), 1—39.

[28]Burke, J. V. , Ferris, M. C. , Qian, M-J. , On the Clarke subdifferential of the distance function of a closed set, J. Math. Anal. Appl. , **166** (1992), 199—213.

[29]Chabrillac, Y. , Crouzeix, J-P. , Continuity and differentiability properties of monotone real functions of several real variables, Math. Prog. Study, **30** (1987), 1—16.

[30]Chaney, R. W. , On second derivatives for nonsmooth functions, Nonlinear Anal. , **9**(1985), 1189—1209.

[31] Chaney. R. W. , Second-order necessary conditions in semismooth optimization, Math. Prog. Study, **40**(1988), 95—109.

[32] Chaney, R. W. , Second-order sufficient conditions in nonsmooth optimization, Math. Oper. Res. , **13**(1988), 660—673.

[33]Chaney, R. W. , Piecewise $C^k$ functions in nonsmooth analysis, Nonlinear Anal. , **15**(1990), 649—660.

[34]陈文塬. 非线性泛函分析. 兰州：甘肃人民出版社, 1982.

[35]Clarke, F. H. , Generalized gradients and applications, Trans. Amer. Math. Soc. , **205**(1975), 247—262.

[36]Clarke, F. H. , Optimization and Nonsmooth Analysis, Wiley, 1983.

[37]Coffman, C. V. , Lyusternik-Schnirelman theory: complementary principles and the Morse index, Nonlinear Anal. , **12**(1988), 507—530.

[38]Conley, C. , Zehnder, E. , Morse type index theory for flows and periodic solutions for Hamiltonian equations, Comm. Pure Appl. Math. **37**(1984), 207—253.

[39] Costa, D. G. , Goncalves, J. V. A. , Critical point theory for nondifferentiable functionals and applications, J. Math. Anal. Appl. , **153** (1990), 470—485.

[40]Dancs, S. , Generalized tangent cone and an optimization problem in a normed space, J. Optim. Theor. Appl. , **67**(1990), 43—55.

[41]Dang, D-H. , Schmitt, K. , Boundary value problems for higher order nonlinear differential equations, Nonlinar Anal. , **21**(1993), 293—305.

[42]Deimling, K. , Ordinary Differential Equations in Banach Spaces, Springer-

Verlag,1977.

[43]Deimling,K. ,Nonlinear Functional Analysis,Springer-Verlag,1985.

[44]Dunn,J. C. ,Tian,T. ,Variants of the Kuhn-Tucker sufficient conditions in cones of nonnegative functions, SIAM J. Control Optim. , **30** (1992), 1361—1384.

[45]Eberhard,A. C. ,Characterization of subgradients.I,J. Math. Anal. Appl. , **132**(1988),57—74.

[46]Ekeland,I. ,On the Variational principle,J. Math. Anal. Appl. ,**47**(1974), 324—353.

[47]Ekeland,I. ,Nonconvex minimization problems,Bull. ,Amer. Math. Soc. , **1**(1979),443—474.

[48] Facchinei, F. , Refinements of necessary conditions for optimality in nonlinear programming,J. Omptim. Theor. Appl. ,**73**(1992),65—74.

[49]Fadell,E. R. ,The relationship between Ljusternik-Schnirelman category and the concept of genus,Pacific J. Math. ,**89**(1980),33—42.

[50]Fischer,Th. ,On the duality of a non-convex optimization problem and the strong unicity constant in linear Chebyshev approximation,J. Math. Anal. Appl. ,**164**(1992),167—177.

[51]Floudas,C. A. ,Visweswaran,V. ,Primal-relaxed dual global optimization approach,J. Optim. Theor. Appl. ,**78**(1993),187—226.

[52]Giannoni,F. ,Periodic solutions of dynamical systems by a saddle point theorem of Rabinowitz,Nonlinear Anal. ,**13**(1989),707—719.

[53] Gomes, S. M. , Sprekels, J. , Krasnoselskii's theorem on operators compressing a cone: Application to some singular boundary value problems,J. Math. Anal. Appl. ,**153**(1990),443—459.

[54]Guddat,J. ,Jongen,H. Th. ,Structural stability in nonlinear optimization, Optimization,**18**(1987),617—631.

[55]郭大钧．非线性泛函分析．济南:山东科学技术出版社,1985.

[56]Guo Dajun,The number of nontrial solutions to Hammerstein nonlinear integral equations,Chin. Ann. of Math. **7B**:2(1986),191—204.

[57]Guo Dajun,Some fixed point theorems and applications,Nonlinear Anal. ,

10(1986),1293—1302.

[58] Guo Dajun, Positive fixed points and eigenvectors of noncompact decreasing operators with applications to nonlinear integral equations, Chin. Ann. of Math. ,**14B**:4(1993),419—426.

[59]Hadjisavras, N. , Kravvaritis, D. , Pantelidis, G. , Polyrakis, I. , Nonlinear monotone operators with value in $\mathscr{L}(X,Y)$, J. Math. Anal. Appl. , **140** (1989),83—94.

[60]Hale, J. , Theory of Functional Differential *Equations*, *Springer*—Verlag, 1977.

[61]Hale, J. , Magalhaes, L. T. , Oliva, W. M. , An Introduction to Infinite Dimensional Dynamical Systems——Geometric Theory, Springer-Verlag, 1984.

[62] Hanson, M. A. , Mond, B. , Necessary and sufficient conditions in constrained optimization, Math. Prog. Study, **37**(1987),51—58.

[63]Hardy, G. H. , Littlewood, J. E. , Pòlya, G. , Inequalities, Cambridge Univ. Press, 1952.

[64]Harker, P. T. , Pang, J-S. , Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications, Math. Prog. Study, **48**(1990),161—220.

[65]Hartman, Ph. , Ordinary Differential Equations, 2ed, Birkhäuser, 1982.

[66]Hekkilä, S. , On the quasimonotonicity of differential systems, Nonlinear Anal. , **7**(1983),91—96.

[67] Hekkilä, S. , Lakshmikantham, V. , Yong Sun, Fixed point results in ordered normed spaces with applications to abstract and differential equations, J. Math. Anal. Appl. , **163**(1992),422—437.

[68]Heinz, H P. , On the behavior of measures of noncompactness with respect to differentiation and integration of vectorvalued functions, Nonlinear Anal. , **7**(1983),1351—1371.

[69] Henrion, R. , On constraint qualifications, J. Optim. Theor. Appl. , **72** (1992),187—197.

[70] Hille, E, Phillips, R. S. , Functional Analysis and Semigroups, Waverly

Press, 1957.

[71] Hiriart-Urruty, J. B., Tangent cones, generalized gradients and mathematical programming in Banach spaces, Math. Oper. Res., 4(1979), 79—97.

[72] Hirsch, M. W., Systems of differential equations that are competitive or cooperative, I: limit sets, SIAM J. Math. Anal., 13(1982), 167—179.

[73] Hirsch, M. W., The dynamical system approach to differential equations, Bull. Amer. Math. Soc., 11(1984), 1—64.

[74] Hirsch, M. W., Systems of differential equations that are competitive or cooperative Ⅱ: convergence almost everywhere, SIAM J. Math. Anal., 16(1985), 423—439.

[75] 胡适耕．现代分析引论．武汉：华中理工大学出版社，1989.

[76] 胡适耕，一类泛函微分方程生成的 K 型单调半流，数学学报，33:6(1990)，820—827.

[77] 胡适耕，一类 K 型单调的积分微分方程，系统科学与数学，11:4(1991)；320—326.

[78] 胡适耕，一类积分微分方程正解的渐近状态，数学年刊，13A(增刊)(1992)，110—115.

[79] Hu Shigeng, Hong Shihuang, A characterization of compact sets in $L^p$ spaces over a locally compact group, in "Applied Functional Analysis Series", Vol. 1, Inter. Acad. Publ., 1993.

[80] Ionescu, I. R., Rosca, I., Friedrichs extensions for nonconvex variational problems, Nonlinear Anal., 14(1990), 905—914.

[81] Irwin, M. C., Smooth Dynamical Systems, Acad. Press, 1980.

[82] Ito, K., Kunisch, K., The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces, Math. Prog. Study, 46(1990), 341—360.

[83] Izé, A. F., On a fixed point index method for the analysis of the asymptotic behavior and boundary value problems of infinite dimensional dynamical systems and processes, J. Diff. Eqs., 52(1984), 162—174.

[84] Jeyakumar, V., Duality and infinite dimensional optimization, Nonlinear

Anal. ,**15**(1990),1111—1122.

[85] Jeyakumar, V., Infinite—dimensional convex programming with applications to constrained approximation, J. Optim. Theor. Appl., **75** (1992),569—586.

[86] Jeyakumar, V., Wolkowicz, H., Zero duality gaps in infinite-dimensional programming, J. Optim. Theor. Appl., **67**(1990),87—108.

[87] Jeyakumar, V., Gwinner, J., Inequality systems and optimization, J. Math. Anal. Appl., **159**(1991),51—71.

[88] Jiang Jifa, Attractors in strongly monotone flows, J. Math. Anal. Appl., **162**(1991),210—222.

[89] Jofre, A., Thibault, L., Proximal and Fréchet normal formulae for some small normal cones in Hilbert spaces, Nonlinear Anal., **19**(1992),599—612.

[90] Jongen, H. T., Twilt, F., Weber, G. W., Semi-infinite optimization: structure and stability of the feasible set, J. Optim. Theor. Appl., **72**(1992),529—552.

[91] Jovanov, D. S., Variational inequalities with cone constraints, J. Optim. Theor. Appl., **75**(1992),87—99.

[92] Kahn, D. W., Introduction to Global Analysis, Acad. Press, 1980.

[93] Karamardian, S., Schaible, S., Seven kinds of monotone maps, J. Optim. Theor. Appl., **66**(1990),37—46.

[94] Karamardian, S. Schaible, S., Crouzeix, J. P., Characterizations of generalized monotone maps, J. Optim. Theor. Appl., **76**(1993),399—413.

[95] Kartsatos, A. G., On the solvability of abstract operator equations involving compact perturbations of m-accretive operators, Nonlinear Anal., **11**(1987),997—1004.

[96] Kawasaki, H., An envelope-like effect of infinitely many inequality constraints on second order necessary conditions for minimization problems, Math. Prog. Study, **41**(1988),73—96.

[97] Kelley, J. General Topology, Van Nostrand, 1975.

[98] Kenderov, P., Multivalued monotone mappings are almost everywhere

single-valued, Studia Math., 56(1976), 199—203.

[99] Khanh, P. Q., Nuong, T. H., On necessary and sufficient conditions in vector optimization, J. Optim. Theor. Appl., 63(1989), 391—413.

[100] Kinderlehrer, D., Stampacchia, G., An Introduction to Variational Inequalities and their Applications, Acad. Press, 1980.

[101] Kotarski, W., On some specification of the Dubovitskii-Milyutin theorem for Pareto optimal problems, Nonlinear Anal., 13(1990), 287—291.

[102] Köthe, G., Topological Vector spaces I, Springer-Verlag, 1983.

[103] Krasnoselskii, M. A., Zabreiko, P. P., Geometrical Methods of Nonlinear Analysis, Springer, 1984.

[104] Krause, U., Ranft, P., A limit set trichotomy for monotone nonlinear dynamical systems, Nonlinear Anal., 19(1992), 375—392.

[105] Krause, U., Nussbaum, R. D., A limit set trichotomy for self-mappings of normal cones in Banach spaces, Nonlinear Anal., 20(1993), 855—870.

[106] Krauss, E., A representation of arbitrary maximal monotone operators via subgradients of skew—symmetric saddle functions, Nonlinear Anal., 9(1985), 1381—1399.

[107] Lang, S., Real Analysis, 2ed. Addison-Wesley, 1983.

[108] Lang. S., Differential Manifolds, Springer-Verlag, 1985.

[109] Lazer, A. C., Solimini, S., Nontrial solutions of operator equations and Morse indices of critical points of minmax type, Nonlinear Anal., 12(1988), 761—775.

[110] Ledzewicz-Kowalewska, U., A necessary and sufficient condition for Altman's maximal element problem for constraint sets with many equalities, Nonlinear Anal., 13(1989), 833—840.

[111] Ledzewicz-Kowalewska, U., Optimality and Pareto optimality conditions for the problems with nonregular operator equality constraints, Nonlinear Anal., 17(1991), 347—360.

[112] Ledzewicz-Kowalewska, U., Euler-Lagrange equation in the case of nonregular equality constraints, J. Optim. Theor. Appl., 71(1991), 549—568.

[113]Lee, J. W., O'Regan, D., Nonlinear boundary value problems in Hilbert spaces, J. Math. Anal. Appl., **137**(1989), 59—69.

[114]Li, T-Y., Yorke, J., Period three implies chaos, Amer. Math. Monthly, **82**(1975), 985—992.

[115]李正元,钱敏,向量场的旋转度及其应用,北京:北京大学出版社,1982.

[116]廖山涛,微分动力系统的定性理论,北京:科学出版社,1992.

[117]Lions, J. L., Stampacchia, G., Variational inequalities, Comm. Pure Appl. Math., **20**(1967), 493—519.

[118]Loewen, P. D., The proximal normal formula in Hilbert space, Nonlinear Anal., **11**(1987), 979—995.

[119]麦结华,连续函数中单峰轨道系列的完整性及 Sarkovskii 定理的推广,中国科学,1989:12,1233—1241.

[120]Martein, L., Lagrange multipliers and generalized differentiable functions in vector extremum problems, J. Optim. Theor. Appl., **63**(1989). 281—297.

[121] Maruyama, Y., Second-order necessary conditions for nonlinear optimization problems in Banach spaces and their application to an optimal control problem, Math. Oper. Res., **15**(1990), 467—482.

[122] Michalek, R., A $Z^p$ Borsuk-Ulam theorem and index theory with a multiplicity result in partial differential equations, Nonlinear Anal., **13**(1989), 957—968.

[123]Mönch, H., Boundary value problems for nonlinear ordinary differential equations of second order in Banach spaces, Nonlinear Anal., **4**(1980), 985—999.

[124] Morales, C. H., On the range of sum of accretive and continuous operators in Banach spaces, Nonlinear Anal., **19**(1992), 1—9.

[125]Moreau, J. J., Valalier, M., A chain rule involving vector functions of bounded variation, J. Funct. Anal., **74**(1987), 333—345.

[126]Munkres, J. R., Elements of Algebraic Topology, Addison-Wesley, 1984.

[127] Naniewicz, Z., On some nonconvex variational problems related to hemivariational inequalities, Nonlinear Anal., **13**(1989), 87—100.

[128] Nemeth, A. B., A nonconvex vector minimization problem, Nonlinear Anal., **10**(1986), 669—678.

[129] Nirenberg, L., Variational and topological methods in nonlinear problems, Bull. Amer. Math. Soc., **4**(1981), 267—302.

[130] Njoku, F. I., Zanolin, F., Positive solutions for two-point boundary value problems: existence and multiplicity results, Nonlinear Anal., **13**(1989), 1329—1338.

[131] Noll, D., A unified approach to duality in convex programming, J. Math. Anal. Appl., **161**(1991), 508—521.

[132] Ntouyas, S. K., Sficas, Y. G., Tsamatos, P. Ch., An existence principle for boundary value problems for second order functional differential equations, Nonlinear Anal., **20**(1993), 215—222.

[133] Palais, R. S., Lusternik-Schnirelman theory on Banach manifolds, Topology, **5**(1966), 115—132.

[134] Palais, R. S., Smale, S., A generalized Morse theory, Bull. Amer. Math. Soc., **70**(1964), 165—171.

[135] Palata, J., A survey of conical approximations used in the optimization, Optimization, **20**(1989), 147—162.

[136] Palis, J., de Melo, W., Geometric Theory of Dynamical Systems, Springer—Verlag, 1982.

[137] Papageorgiou, N. S., Nonsmooth analysis on partially ordered vector spaces, Part 1: convex case, Pacif. J. Math. **107**(1983), 403—458.

[138] Papageorgiou, N. S., Nonsmooth analysis on partially ordered vector spaces, Part 2: nonconvex case, Clarke's theory, Pacif. J. Math., **109**(1983), 463—495.

[139] Papageorgiou, N. S., Nonsmooth analysis on partially ordered vector spaces: the subdifferential theory, Nonlinear Anal., **10**(1986), 615—637.

[140] Pellegrini, L., On a Lagrangian sufficient optimality condition, J. Optim. Theor. Appl., **68**(1991), 19—33.

[141] Penot, J-P., A characterization of tangential regularity, Nonlinear Anal., **5**(1981), 625—663.

[142] Penot, J-P., The drop theorem, the petal theorem and Ekeland's variational principle, Nonlinear Anal., **10**(1986), 813—822.

[143] Poliquin, R. A., Subgradient monotonicity and convex functions, Nonlinear Anal., **14**(1990), 305—317.

[144] Pounder, J. R., Rogers, D., Dynamics of a two-parameter family of maps of the interval, Nonlinear Anal., **10**(1986), 415—423.

[145] Pucci, P., A mountain pass theorem, J. Diff. Eqs., **10**(1986), 615—637.

[146] Pucci, P., Serrin, J., Extensions of the mountain pass theorem, J. Funct. Anal., **59**(1984), 185—210.

[147] Rabier, P. J., Topological degree and the theorem of Borsuk for general covariant mappings with applications, Nonlinear Anal., **16**(1991), 399—420.

[148] Reiland, T. W., A gemetric approach to nonsmooth optimization with sample applications, Nonlinear Anal., **10**(1987), 1169—1184.

[149] Revalski, J. P., Zhivkov, N. V., Well-posed constrained optimization problems in metric spaces, J. Optim. Theor. Appl., **76**(1993), 145—163.

[150] Robinson, S. M., An implicit-function theorem for a class of nonsmooth functions, Math. Oper. Res., **16**(1991), 292—309.

[151] Rockafellar, R. T., On the maximal monotonicity of subdifferentiable mappings, Pacif. J. Math., **33**(1970), 209—216.

[152] Rockafellar, R. T., Convex Analysis, 2ed. Univ. Press, 1972.

[153] Rockafellar, R. T., Directionally Lipschitzian functions and subdifferential calculus, Proc. London Math. Soc., **39**(1979), 331—335.

[154] Rockafellar, R. T., Generalized directional derivatives and subgradients of nonconvex functions, Canad. J. Math., **32**(1980), 257—280.

[155] Rockafellar, R. T., Lipschitzian properties of multifunctions, Nonlinear Anal., **9**(1985), 867—885.

[156] Rockafellar, R. T., Extensions of subgradient calculus with applications to optimization, Nonlinear Anal., **9**(1985), 665—698.

[157] Rothe, E., On the connection between critical point theory and Leray Schauder degree, J. Math. Anal. Appl., **88**(1982), 265—269.

[158]Santanilla, J., Some coincidence theorems in wedges, cones and convex sets, J. Math. Anal. Appl., **105**(1985), 357—371.

[159]Santanilla, J., Nonnegative solutions of boundary value problems for nonlinear first and second order ordinary differential equations, J. Math. Anal. Appl., **126**(1987), 397—408.

[160]Schmitt, K., Smith, H. L., On eigenvalue problems for nondifferentiable mappings, J. Diff. Eqs., **33**(1979), 294—319.

[161]Schwartz, J. T., Generalizing the Lusternik-Schnirelman theory of critical points, Comm. Pure Appl. Math., **17**(1964), 307—315.

[162]史树中，非光滑分析，数学进展，**15**:1(1986), 9—21.

[163]Chow, Shui-Nee, Lu, K., Invariant manifolds for flows in Banach spaces, J. Diff. Eqs., **74**(1988), 285—317.

[164]Siegberg, H. W., Some historycal remarks concerning degree theory, Amer. Math. Monthly, **88**(1981), 125—139.

[165]Singer, I., A general theory of dual optimization problems, J. Math. Anal. Appl., **116**(1986), 77.

[166]Singer, I, Some general Lagrangian duality theorems, J. Math. Anal. Appl. **144**(1989), 26—51.

[167]Singer, I., Some futher duality theorems for optimization problems with reverse convex constraint sets, J. Math. Anal. Appl., **171**(1992), 205—219.

[168]Smale, S., Differential dynamical systems, Bull. Amer. Math. Soc., **73**(1967), 747—817.

[169]Smith, H., Monotone semiflows generated by functional differential equations, J. Diff. Eqs., **66**(1987), 420—442.

[170]Smith, H., Systems of ordinary differential equations which generate an order preserving flow, SIAM Rev., **30**(1988), 87—113.

[171]Smith, H., Thieme, H. R., Monotone semisflows in scalar nonquasi-monotone functional differential equations, J. Math. Anal. Appl., **150**(1990), 289—306.

[172]Smith, H., Thieme, H. R. Convergence for strongly order-preserving

semiflows, SIAM J. Math. Anal., 22(1991), 1081—1101.

[173]Smith, H., Thieme, H. R., Strongly order preserving semiflows generated by fde, J. Diff. Eqs., 93(1991), 332—363.

[174]Studniarski, M., Mean value theorems and sufficient optimality conditions for nonsmooth functions, J. Math. Anal. Appl., 111(1985), 313—326.

[175] Sun Jingxian, Fixed point theorems of increasing operators and applications to nonlinear integro-differential equations with discontinuous terms, J. Math. Anal. Appl., 175(1993), 33—45.

[176]孙经先，两点拉伸型不动点定理及其应用，系统科学与数学，12(1992)，284—286.

[177] Szufla, Boundary value problems for nonlinear ordinary differential equations of second order in Banach spaces, Nonlinear Anal., 8(1984), 1481—1487.

[178]Szulkin, A., A relative category and applications to critical point theory for strongly indefinite functionals, Nonlinear Anal., 15(1990), 725—739.

[179] Takáč, P., Asymptotic behavior of discrete-time semigroups of sublinear, strongly increasing mappings with applications to biology, Nonlinear Anal., 14(1990), 35—42.

[180] Tersian, S. A., A minimax theorem and applications to nonresonance problems for semilinear equations, Nonlinear Anal., 10(1986), 651—668.

[181]Thach, P. T., Convex minimization under Lipschitz constraints, J. Optim. Theor. Appl., 64(1990), 595—614.

[182] Thibault, L., Subdifferentials of nonconvex vector-valued functions, J. Math. Anal. Appl., 86(1982), 319—344.

[183]Treiman J. S., Characterization of Clarke's tangent and normal cones in finite and infinite dimension, Nonlinear Anal., 7(1983), 777—783.

[184]Treiman, J. S., Clarke's gradients and epsilon-subgradients in Banach spaces, Trans. Amer. Math. Soc., 294(1986), 65—78.

[185] Treiman, J. S., Shrinking generalized gradients, Nonlinear Anal., 12 (1988), 1429—1450.

[186]Treiman, J. S., An infinite class of convex tangent cones, J. Optim.

Theor. Appl. ,**68**(1991),563—581.

[187]Uhlenbech,K. ,Morse theory on Banach manifolds,J. Funct. Anal. ,**10** (1972),430—445.

[188]Vidossich,G. ,On the solvability of boundary value problems for higher order ordinary differential equations, Nonlinear Anal. ,**13**(1989),1171—1180.

[189]Vidossich,G. ,A general existence theorem for boundary value problems for ordinary differential equations,Nonlinear Anal. ,**15**(1990),897—914.

[190]Wang,T-X. ,Lusternik-Schnirelman category theory on closed subsets of Banach manifolds,J. Math. Anal. Appl. ,**149**(1990),412—423.

[191] Ward, D., Isotone tangent cones and nonsmooth optimization, Optimization,**18**(1987),769—783.

[192] Ward D. , Which subgradients have sum formulas, Nonlinear Anal., **11**(1988),1231—1243.

[193]Ward,D. ,Chain rules for nonsmooth functions,J. Math. Anal. Appl., **158**(1991),519—538.

[194]Weckesser,V. ,The subdifferential in Banach spaces,Nonlinear Anal., **20**(1993),1349—1354.

[195]Weinstein,S. E. ,Xu Yusheng,A duality approach to best uniform convex approximation,J. Math. Anal. Appl. ,**160**(1991),314—322.

[196] Yang Guang-Hong, The generalized topological degree for the perturbations of m-accretive operators and its applications, Nonlinear Anal. ,**18**(1992),403—412.

[197]Yang Xuefeng,The Morse critical groups of the minimax theorem,Chin. Ann. of Math. ,**12B**;2(1991),202—206.

[198]Yosida,K. ,Functional Analysis ,Springer-Verlag,1978.

[199]Yu Qingyu,Cheng Jiangang,A bifurcation theorem of set—contractive maps,Chin. Ann. of Math. ,**6B**;2(1985),251—255.

[200]Yuan Yaxiang,Nonlinear programming—— a review,Adv. Math. ,**19**;1 (1990),12—24.

[201]Zeidler, E. , The Lusternik-Schnirelman theory for indefinite and not

necessarily odd nonlinear operators and its applications, Nonlinear Anal., **4**(1980), 451—489.

[202]Zeidler, E., Nonlinear Functional Analysis and its Applications, I, II, III Springer-Verlag, 1986.

[203]张恭庆,一个变化的 Pountain Pass 定理,中国科学,A.(1983),306—317.

[204]张恭庆,临界点理论及其应用.上海:上海科学技术出版社,1986.

[205]张景中,杨路,关于 Sarkovskii 序的一些定理,数学进展,**16**:1(1987),33—48.

[206]张锦炎,钱敏,微分动力系统导引,北京:北京大学出版社,1991.

[207]张石生,不动点理论及应用,重庆:重庆出版社,1984.

[208]张石生,陈玉清,增生映射方程研究的重合指数方法,数学年刊,**14A**:5(1993),579—583.

[209]张筑生,微分动力系统原理,北京:科学出版社,1987.

[210]Zhao Weiyu, Remarks on various measures of noncompactness, J. Math. Anal. Appl., **174**(1993), 290—297.

[211]Zho Yichun, On a surjectivity for the sum of two mappings of monotone type, Chin. Ann. of Math. **6B**(1985), 471—480.

[212]赵义纯,非线性泛函分析及其应用,北京:高等教育出版社,1989.

[213]钟承奎,Banach 空间中等变全连续场的拓扑度计算,数学年刊,**12A**:2(1991),154—159.

[214]周叔子,变分不等式及其有限元方法,长沙:湖南大学出版社,1988.

necessarily odd nonlinear operators and its applications, Nonlinear Anal., **4**(1980), 451—489.

[202]Zeidler, E., Nonlinear Functional Analysis and its Applications, I, II, III Springer-Verlag, 1986.

[203]张恭庆,一个变化的 Pountain Pass 定理,中国科学,A. (1983), 306—317.

[204]张恭庆,临界点理论及其应用.上海:上海科学技术出版社,1986.

[205]张景中,杨路,关于 Sarkovskii 序的一些定理,数学进展,**16**:1(1987), 33—48.

[206]张锦炎,钱敏,微分动力系统导引,北京:北京大学出版社,1991.

[207]张石生,不动点理论及应用,重庆:重庆出版社,1984.

[208]张石生,陈玉清,增生映射方程研究的重合指数方法,数学年刊,**14A**:5(1993), 579—583.

[209]张筑生,微分动力系统原理,北京:科学出版社,1987.

[210]Zhao Weiyu, Remarks on various measures of noncompactness, J. Math. Anal. Appl., **174**(1993), 290—297.

[211]Zho Yichun, On a surjectivity for the sum of two mappings of monotone type, Chin. Ann. of Math. **6B**(1985), 471—480.

[212]赵义纯,非线性泛函分析及其应用,北京:高等教育出版社,1989.

[213]钟承奎,Banach 空间中等变全连续场的拓扑度计算,数学年刊,**12A**:2(1991), 154—159.

[214]周叔子,变分不等式及其有限元方法,长沙:湖南大学出版社,1988.

necessarily odd nonlinear operators and its applications, Nonlinear Anal., **4**(1980), 451—489.

[202]Zeidler, E., Nonlinear Functional Analysis and its Applications, Ⅰ, Ⅱ Ⅲ Springer-Verlag, 1986.

[203]张恭庆，一个变化的 Pountain Pass 定理，中国科学，A.(1983)，306—317.

[204]张恭庆，临界点理论及其应用.上海：上海科学技术出版社，1986.

[205]张景中，杨路，关于 Sarkovskii 序的一些定理，数学进展，**16**：1(1987)，33—48.

[206]张锦炎，钱敏，微分动力系统导引，北京：北京大学出版社，1991.

[207]张石生，不动点理论及应用，重庆：重庆出版社，1984.

[208]张石生，陈玉清，增生映射方程研究的重合指数方法，数学年刊，**14A**：5(1993)，579—583.

[209]张筑生，微分动力系统原理，北京：科学出版社，1987.

[210]Zhao Weiyu, Remarks on various measures of noncompactness, J. Math. Anal. Appl. **174**(1993), 290—297.

[211]Zho Yichun, On a surjectivity for the sum of two mappings of monotone type, Chin. Ann. of Math. **6B**(1985), 471—480.

[212]赵义纯，非线性泛函分析及其应用，北京：高等教育出版社，1989.

[213]钟承奎，Banach 空间中等变全连续场的拓扑度计算，数学年刊，**12A**：2(1991)，154—159.

[214]周叔子，变分不等式及其有限元方法，长沙：湖南大学出版社，1988.

necessarily odd nonlinear operators and its applications, Nonlinear Anal.,4(1980),451—489.

[202]Zeidler,E.,Nonlinear Functional Analysis and its Applications, I, II, III Springer-Verlag,1986.

[203]张恭庆,一个变化的 Pountain Pass 定理,中国科学,A.(1983),306—317.

[204]张恭庆,临界点理论及其应用.上海:上海科学技术出版社,1986.

[205]张景中,杨路,关于 Sarkovskii 序的一些定理,数学进展,**16**:1(1987),33—48.

[206]张锦炎,钱敏,微分动力系统导引,北京:北京大学出版社,1991.

[207]张石生,不动点理论及应用,重庆:重庆出版社,1984.

[208]张石生,陈玉清,增生映射方程研究的重合指数方法,数学年刊,**14A**:5(1993),579—583.

[209]张筑生,微分动力系统原理,北京:科学出版社,1987.

[210]Zhao Weiyu,Remarks on various measures of noncompactness,J. Math. Anal. Appl.,**174**(1993),290—297.

[211]Zho Yichun,On a surjectivity for the sum of two mappings of monotone type,Chin. Ann. of Math. **6B**(1985),471—480.

[212]赵义纯,非线性泛函分析及其应用,北京:高等教育出版社,1989.

[213]钟承奎,Banach 空间中等变全连续场的拓扑度计算,数学年刊,**12A**:2(1991),154—159.

[214]周叔子,变分不等式及其有限元方法,长沙:湖南大学出版社,1988.